SNOMED International

Leading healthcare terminology, worldwide

Practical Guide to Reference Sets

31 March 2017

Latest web browsable version: http://snomed.org/refsetpg

SNOMED CT Document Library: http://snomed.org/doc

Practical Guide to Reference Sets

SNOMED International — Leading healthcare terminology, worldwide

The SNOMED CT Practical Guide to Reference Sets starts by identifying practical use cases and the requirements that must be met to address them. It then explains how different types of reference set can be used to meet those requirement. It also provides advice on different approaches to creating, editing, maintaining and using reference sets to support effective localization and customization of SNOMED CT.

Latest web browsable version: http://snomed.org/refsetpg

SNOMED CT Document Library: http://snomed.org/doc
Publication date: 2017-03-30
This is the date when this PDF file was generated from the online version of this document.

1. Introduction

Background

SNOMED CT provides the core clinical terminology for the electronic health record (EHR) and contains more than 300,000 active concepts with unique meanings. These concepts are organized into hierarchies and have formal logic-based definitions. When implemented in software applications, SNOMED CT can be used to represent clinically relevant information consistently, reliably and comprehensively as an integral part of producing electronic health records. Due to the comprehensiveness and expressivity of SNOMED CT it is often useful to constrain its use to a subset of concepts, descriptions or relationships relevant to a particular use case. SNOMED CT reference sets provide a standard way to represent subsets of SNOMED CT components. Reference sets also provide an extensible mechanism to customize the terminology to meet a wide range of practical requirements.

Purpose

The aim of this document is to provide a high level introduction to SNOMED CT reference sets, and to explain the different types of reference sets and their usage. Furthermore, the document includes an introduction to the reference set format and provides guidance on the development and management of reference sets. Thus, the objective of this document is to support users of SNOMED CT in:

- Understanding the purpose of reference sets
- Knowing about the different types of reference sets and their characteristics
- Choosing the correct type of reference set for a specific purpose
- Creating, developing and maintaining reference sets
- Using reference sets together with other SNOMED CT resources (for entry and display, analytics, knowledge linkage, communication)
- Sharing reference sets
- Adopting or adapting existing reference sets
- Exploring and assessing the content of existing reference sets

Audience

The intended audiences for this guide are those involved in the creation, maintenance and usage of SNOMED CT reference sets. More specifically, this includes:

- SNOMED International Members who wish to learn about the practical uses of reference sets or who are involved with defining reference sets
- Clinicians, informatics specialists and technical staff involved in the planning, management, design or implementation of reference sets.
- Software vendors, data analysts, epidemiologists and others designing SNOMED CT based solutions.

This document assumes a basic level of understanding of SNOMED CT. For background information the reader should refer to the SNOMED CT Starter Guide.

Reference Sets at a Glance

SNOMED CT REFERENCE SETS AT A GLANCE
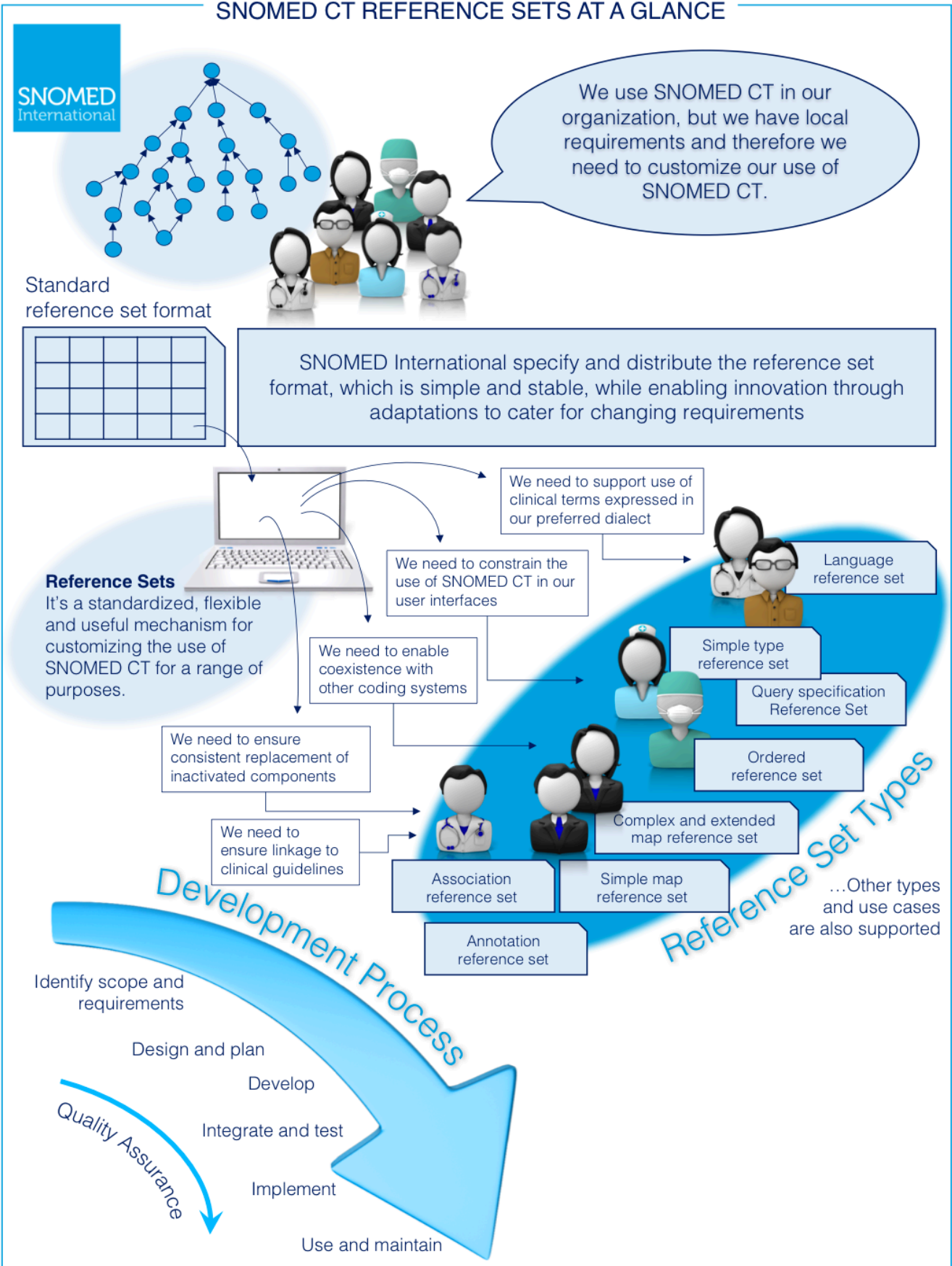
2. Subsets, Value Sets and Reference Sets

In this section, we define some important terms that are used throughout this guide. In particular, subset, value set and reference set.

Subset and value set are general terms that are not specific to SNOMED CT. However, it is important to understand what they mean, and how they relate to SNOMED CT reference sets.

In summary:

- A subset is a set of members all of which are also members of another set.
- A value set is a uniquely identifiable set of valid concept representations, where any concept representation can be tested to determine whether or not it is a member of the value set.
- A reference set is a standard format for maintaining and distributing a set of references to SNOMED CT components and optionally associating referenced components with additional information.

In the following subsections, we describe each of these terms in more detail.

- 2.1. Subset
- 2.2. Value Set
- 2.3. Reference Set

2.1. Subset

A subset is defined as a set of members all of which are also members of another set.

Notes

The definition of subset stated above matches the general use of the word subset in set theory and mathematics. The notes below apply this definition to subsets of SNOMED CT components.

1. A subset of SNOMED CT concepts is a defined set of concepts taken from a wider set of concepts (e.g. all the concepts in a particular version of a specified SNOMED CT Edition).
2. Similarly, a subset of SNOMED CT descriptions is a set of descriptions taken from a wider set of descriptions (e.g. all the descriptions in a particular version of a specified SNOMED CT Edition).
3. The members of a subset can defined in one of two ways extensionally, by enumeration, or intensionally, using rules to determine inclusion.
4. The standard distribution format for extensionally defined subsets is a simple reference set, while the standard distribution format for intensionally defined subsets is query reference sets.

Subset Example

The diagram below shows an example of a subset. The English vowels are a subset of the set of alphabet characters included in the English alphabet
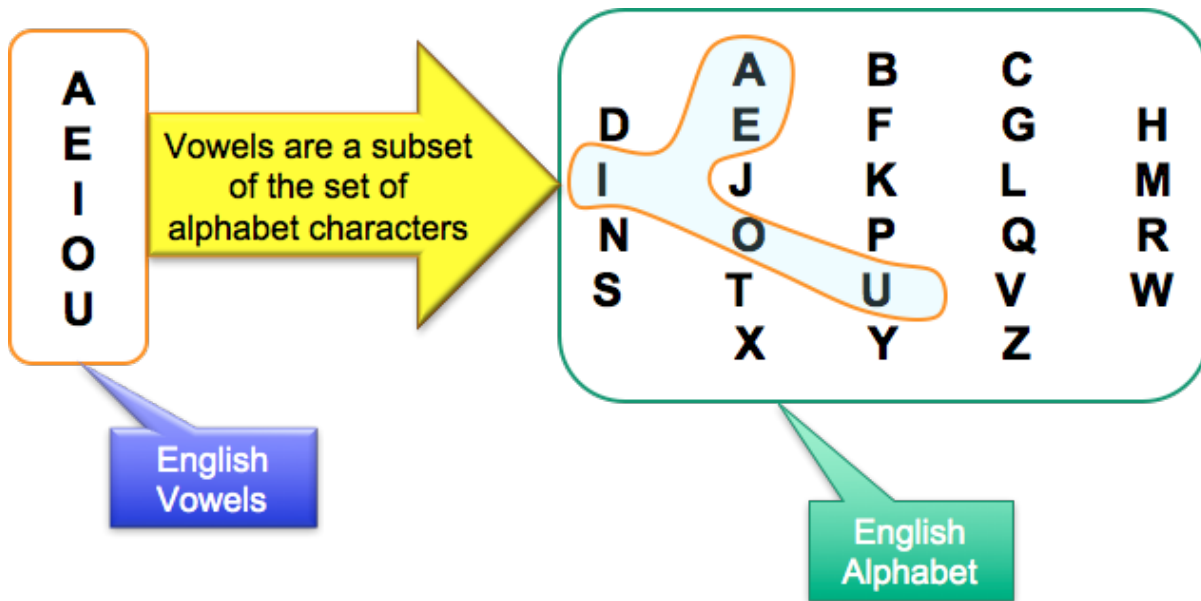
Figure 2.1-1: Subset example

2.1.1. Subset Definitions

There are two distinct ways to define the membership of a subset. These are known as extensional and intensional subset definitions. These terms are defined and illustrated below.

Extensional subset definition
A subset definition in which the membership is represented by enumeration.
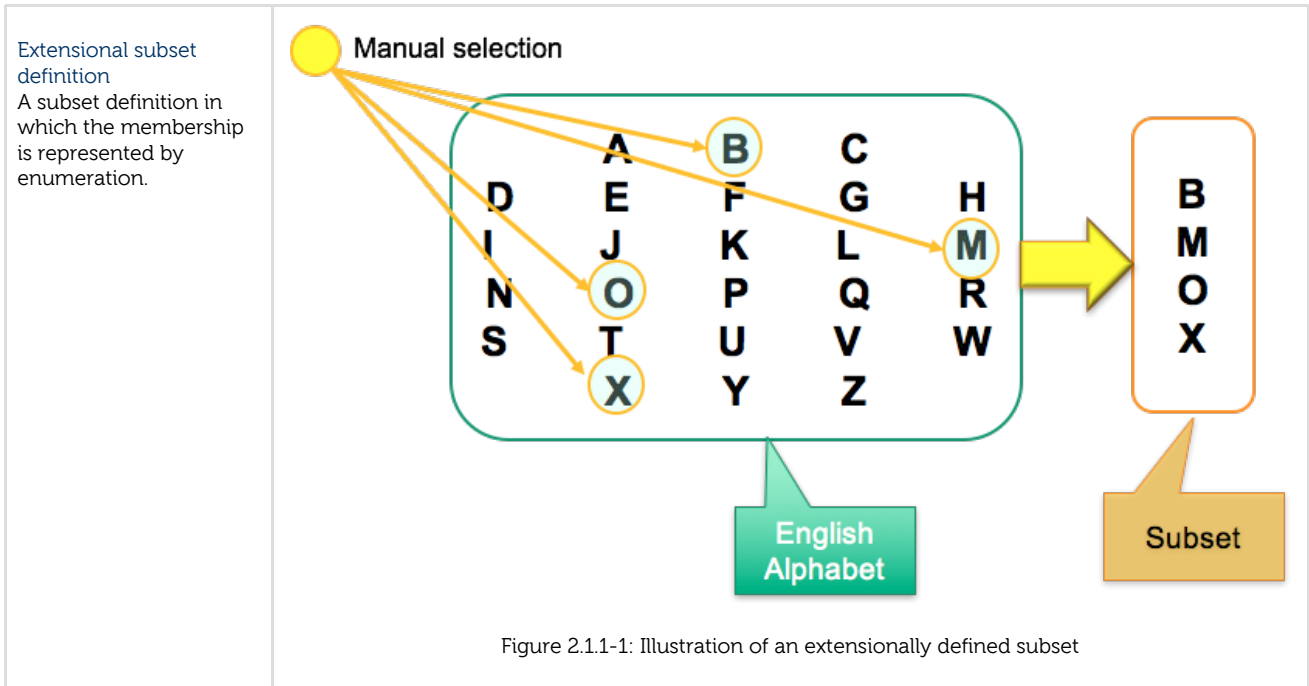
Figure 2.1.1-1: Illustration of an extensionally defined subset

Intensional subset definition
A subset definition in which the membership is represented by a set of rules. These rules specify the necessary conditions that must be met for inclusion within the subset.
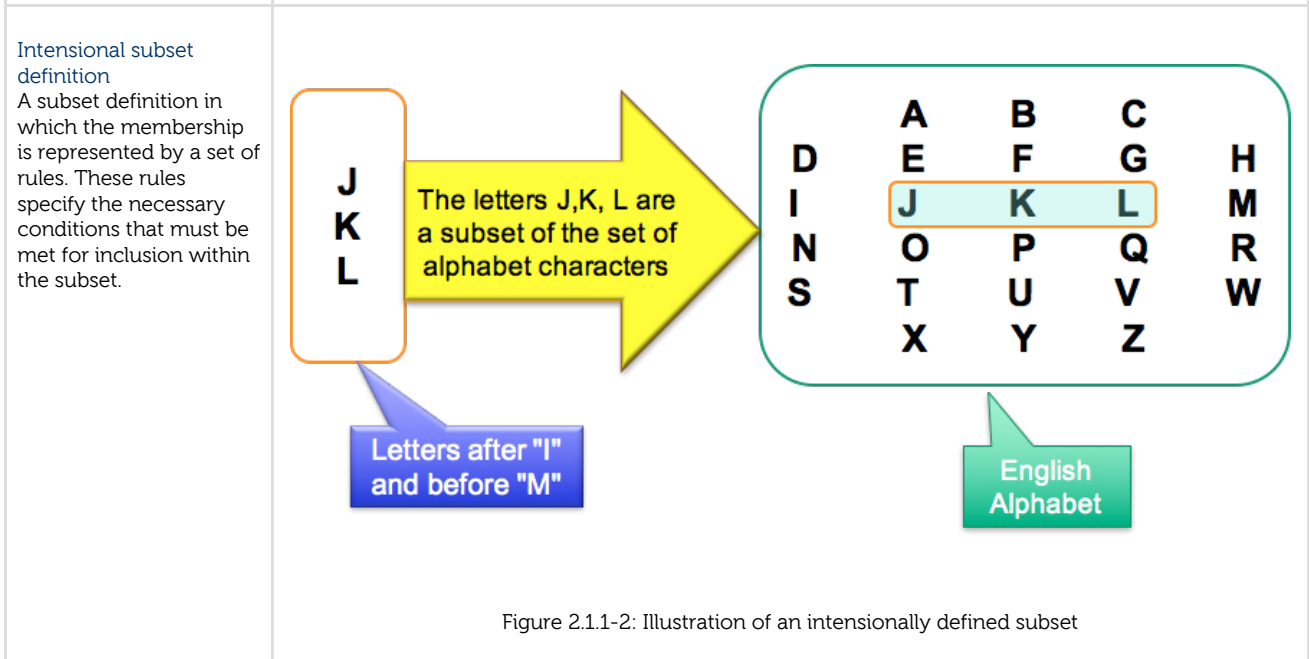
Figure 2.1.1-2: Illustration of an intensionally defined subset

Extensional Definitions

Extensionally defined subsets of SNOMED CT components consist of a list of SNOMED CT component identifiers, which are typically either concept or description identifiers. Extensionally defined SNOMED CT subsets are formally represented using the simple reference set type.
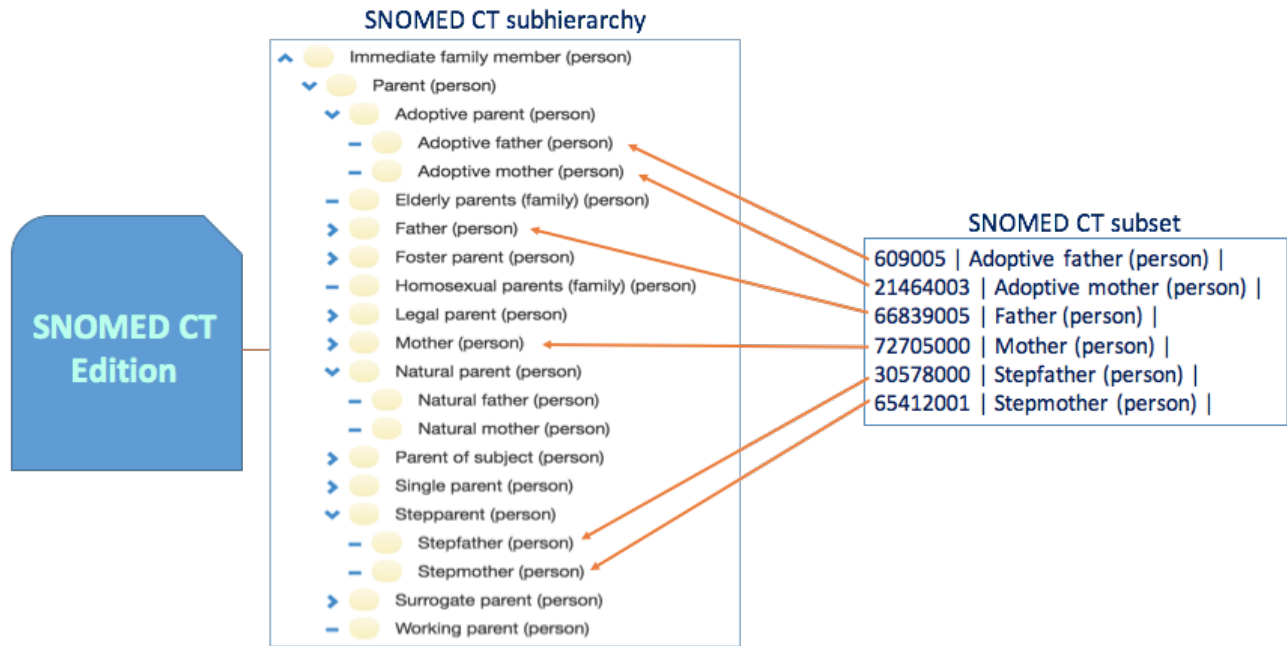
Figure 2.1.1-3: Illustration of an extensionally defined subset of SNOMED CT concepts

Intensional Definitions

The design of SNOMED CT lends itself to intensional subset definitions, because rules can be formulated to specify the conditions for member inclusion. The subtype hierarchy and formal definitions of SNOMED CT concepts enable constraints to be specified, such as:

- Concepts that are descendants of a particular concept.
  - For example: < 118794001 |Procedure on lung| (i.e. all descendants of the concept 118794001 |Procedure on lung|)
- Concepts that share a set of defining characteristics.
  - For example: < 64572001 |Disorder| : 116676008 |Associated morphology| = 257552002 |Inflammation| (i.e. all disorders with an associated morphology of inflammation)
- Concepts that are members of a particular subset AND share one or more defining characteristics.
  - For example: ^ 450984003 |Symptoms and signs reference set for GP/FP health issue| AND < 95324001 |Skin lesion| (i.e. all members of the 450984003 |Symptoms and signs reference set for GP/FP health issue| which are also descendants of 95324001 |Skin lesion|)

Intensionally defined SNOMED CT subsets are formally represented using a query specification reference set . The query string in this reference set represents the intensional definition of the subset. The standard way of representing the query is using the SNOMED CT Expression Constraint Language (ECL).

Substrate

The substrate is the superset of members to which an intensional subset definition is applied. Related to subsets of SNOMED CT components the 'Substrate' is the SNOMED CT content against which an intensional query is executed. Typically, the substrate is a specified version of a particular SNOMED CT Edition. Regardless of whether the subset members are intentionally defined or extensionally defined it is important to be specific about the substrate used to generate the subset, because the result of running a query, or manually selecting the subset members, may vary depending on the substrate used.
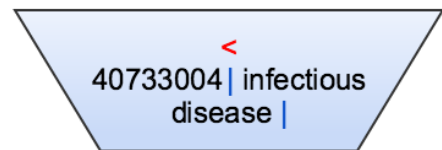
Expansion

The expansion is the result of applying an intensional definition to a given substrate. The Expansion may differ depending on the substrate that the query is run against. The expansion that results from running the query against a specific SNOMED CT substrate can be represented extensionally as a simple reference set. Alternatively, the expansion may be computed dynamically or represented in an internal format within a software application.

For more information about the process and methods for developing subsets and reference sets, see the section about refset development.
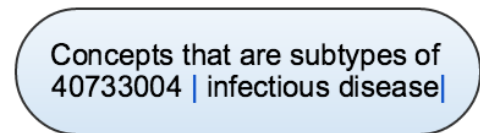
Figure 2.1.1-4: Illustration of a substrate, intentional subset definition and expansion

2.2. Value Set

A value set is a uniquely identifiable set of valid concept representations, where any concept representation can be tested to determine whether or not it is a member of the value set.
A value set is typically used to represent the possible values of a coded data element in an information model. The members of a value set may represent concepts using either simple codes or postcoordinated expressions.

There are a number of use cases for value sets, including constraining the permitted values for elements in a communication specification, specifying the values in a pick list on a user interface and defining the required values to use for reporting. Value sets may range from a simple flat list of codes from a single code system, to an unbounded hierarchical set of post-coordinated expressions drawn from multiple code systems. Value sets containing only SNOMED CT components may be represented as SNOMED CT reference sets.

For example, a message or reporting specification might define a single value set for a problem list, which includes:

- SNOMED CT |Disorder| concepts
- SNOMED CT expressions that are subtypes of |Disorder|
- ICD-10 classification codes representing diseases

Value Set Example

The diagram below illustrates an example of an Observation model, which may be used to support diagnosis, monitor progress, determine patterns in clinical data, etc. Each data element in the information model is linked to a value set, which represents the value values for that element. As shown below, the value sets used in this information model may be selected from different code systems. In some cases, a single value set may also include concepts from different code systems.
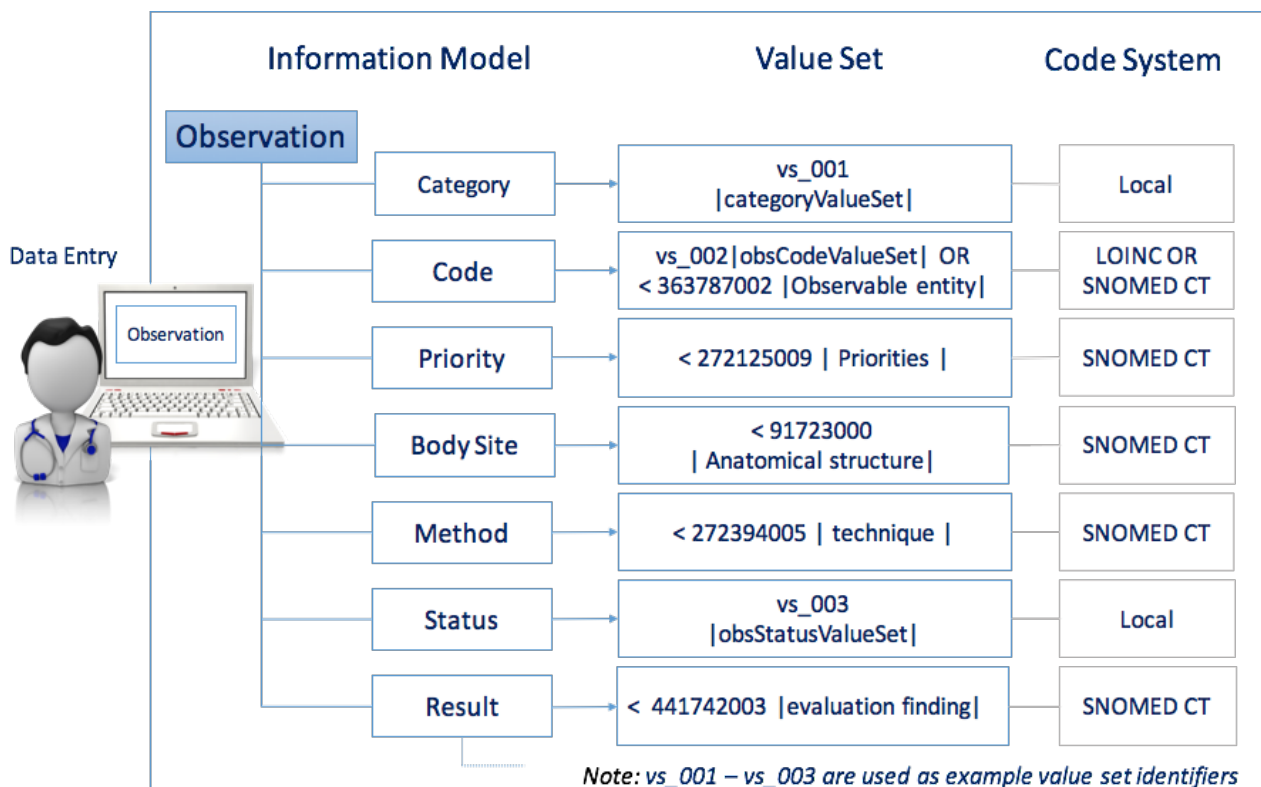
Figure 2.2-1: Value sets used in an information model

2.3. Reference Set

A reference set is defined as: a standard format for maintaining and distributing a set of references to SNOMED CT components and optionally associating referenced components with additional information.

Note

1. A reference set can be used to represent a subset of components (concepts, descriptions or relationships).
2. A reference set that associates additional information with referenced components can be used to support various different purposes, such as representing:
   - An ordered lists of components;
   - Sets of associations between components;
   - Maps between SNOMED CT concepts and codes in other code systems, classifications or knowledge resources.

Hence, reference sets are a mechanism that can be used to represent subsets and value sets of SNOMED CT components. However, they can also be used for many other purposes, including those summarized in 3.2. Use Cases. The following table lists the attributes used in the standard reference set format, and their purpose.

Table 2.3-1: Standard reference set format

Attribute	Purpose
id	General versioning information
effectiveTime	
active	
moduleId	
refsetId	Reference set identifier
referencedComponentId	Component reference
<attribute-1 ... attribute-n>	Additional information depending on reference set type

Reference Sets in an Extension

SNOMED CT is designed to allow the International Edition to be extended to meet national or local requirements. A SNOMED CT Extension may contain components of various types including concepts, descriptions, relationships and derivatives including reference sets. Since the International Release and extensions share a common structure, the same application software can be used to enter, store and process information from different extensions. Extensions can also be shared without requiring additional software procurement or development. Most Extensions are managed by Members or Affiliates of SNOMED International.

Reference sets authored in an extension use the same common format as those in the International Edition. This makes it easier to use the same software to create, maintain and share extension reference sets, as used by international reference sets. The reference set format can also be customized in an extension using a standardized customization approach, as explained in section 4.3. Pre-defined and Customized Reference Sets.

3. Requirements and Use Cases

Reference sets are useful for a range of purposes, and different types of reference sets have been defined to meet various requirements. It is therefore important to understand the requirements for reference set usage in order to be able to select the appropriate reference set design.

Reference sets are useful for including, excluding and prioritizing content or managing the use of codes for data entry, communication or analysis purposes. Reference sets are important when customizing the use of SNOMED CT to meet specific requirements. For example, to meet national, jurisdictional or organizational requirements for recording clinical information, or to assess regional variations in disease prevalence or health delivery.

Practical use cases for SNOMED CT span the entire clinical information lifecycle from data entry, through to storage, display, communication, retrieval, analysis and reuse. Reference sets support effective and efficient use of SNOMED CT for a range of purposes at different stages in that lifecycle. For example, reference sets can assist with user interface customization, specification of reporting criteria, mapping to statistical classifications and representation of links to knowledge resources. They can also be used to represent terminology bindings and to validate the content of both inbound and outbound communications.

The subsequent sections will introduce some of the typical requirements and use cases for reference sets and the different types of reference sets which support these use cases.

- 3.1. Requirements
- 3.2. Use Cases

3.1. Requirements

To decide if a reference set is needed, successful implementers of SNOMED CT must clearly understand their requirements.

The table below shows some typical requirements that may be met using SNOMED CT reference sets. The examples in this table are not exhaustive, but instead illustrate common use cases in which the category of requirement is needed. The table also shows the type (or types) of reference set that best meets each requirement. For more information on each requirement category, please click on the diagram to visit the relevant section.

Table 3.1-1: Requirements met by reference sets

REQUIREMENTS	EXAMPLE USE CASES	REFERENCE SET
3.1.1. A Subset of Components	• 3.2.6.1. Constrain Value Sets • 3.2.3.2. Specifying Queries for Retrieval and Analysis • 3.2.5.2. Interface Terminology	5.1. Simple Reference Set
3.1.2. An Ordered List of Components	• 3.2.1.4. Order Items for Search and Data Entry • 3.2.1.5. Alternative Hierarchical View	5.3. Ordered Reference Set
3.1.3. A Set of Associations between Components	• 3.2.6.3.2. Representing Historical Associations • 3.2.1.6. Use Case Specific Associations	5.4. Association Reference Set
3.1.4. A Set of Components Annotated with Additional Information	• 3.2.2.1. Linking Concepts to Web Resources • Link components to a textual advice	5.5. Annotation Reference Set
3.1.5. A Set of Maps between SNOMED CT and Another Code System	• 3.2.3.1. Maps to Statistical Classifications • Link concepts to legacy codes and data, to support migration	4.2.8. Simple Map Reference Set 4.2.9. Complex and Extended Map Reference Sets
3.1.6. A Set of Sets of Components	• 3.2.6.2. Managing Value Sets • 3.2.6.2. Managing Value Sets	5.1. Simple Reference Set 5.2. Query Specification Reference Set

3.1.1. A Subset of Components

When implementing SNOMED CT in a clinical software system, such as an EHR, the use of SNOMED CT will typically involve customization. For example, selecting a subset of SNOMED CT components to be used for a particular purpose is a typical way of customizing SNOMED CT for use. Subsets of SNOMED CT components can be used to constrain the use of SNOMED CT by either including, excluding or prioritizing specific components. An extensionally defined subset of components, whether it is a subset of concepts, descriptions or relationships, can be represented as a simple reference set. The definition of an intensionally defined subset of components can be represented using a query specification reference set , while its expansion can be represented using a simple reference set. A subset of components may be required to support a range of different uses, as illustrated in the table below.

Table 3.1.1-1: Requirements for subsets of components

Requirement	Description	Example uses	Reference Set
Subset of concepts	An extensionally defined set of references to SNOMED CT concepts	• Restricting searches to terms associated with specified concepts • Constraining data entry • Specifying value sets for particular data items • Specifying queries for data retrieval	Simple reference set
	An intensionally defined set of references to SNOMED CT concepts		Query specification reference set
Subset of descriptions	A set of references to SNOMED CT descriptions	• Restricting searches to specified sets of terms • Specifying descriptions to appear in a list of options	Simple reference set
Inclusion/Exclusion of content	A set which contains the components to be included/ excluded	• Excluding particular components from search and/or data entry • Including a subset of concepts/descriptions for search, data entry, reporting etc.	Simple reference set

For more detailed use case examples, please refer to the following sections:

- Constrain data entry
- Constrain searches
- Exclude content
- Categorising patients using subsets
- Constrain value sets

3.1.2. An Ordered List of Components

Organizing members of a subset into a specific order can be useful to meet certain implementation requirements, such as displaying drop down lists for data entry or search results. Members of a subset can be ordered in a variety of automated ways, including displaying the shortest term that matches the search term first, alphabetically, or randomly. However, when the required order cannot be automatically computed, it may be necessary to specify an order for each subset member. Ordered lists of SNOMED CT components (typically concepts or descriptions) can be represented using an Ordered Reference Set. For more information about ordering SNOMED CT components on a user interface, please refer to the Search and Data Entry Guide.

Table 3.1.2-1: Requirements for an ordered list of components

Requirement	Description	Example Uses	Reference Set
An ordered list of concepts	A subset consisting of references to specific SNOMED CT concepts, i.e. a set of SNOMED CT Identifiers. Additionally, each subset member is assigned a specific order, which enables prioritization.	• Presenting concepts in an order that is rational or helpful for a particular purpose irrespective of the term displayed • Making it easier to find concepts that are most commonly used in a particular specialty, department or data entry scenario	5.3. Ordered Reference Set
An ordered list of descriptions	A subset consisting of references to specific SNOMED CT descriptions, i.e. a set of SNOMED CT Identifiers, where each included member identifies a description. Additionally, each subset member is assigned a specific order, to enable ordering or prioritizing the members.	• Presenting terms in an order that is rational or helpful for a particular purpose in user interface controls including: 　• Simple lists 　• Drop down lists 　• Popup menus • Ordering search results	5.3. Ordered Reference Set

For more detailed use case examples, please refer to the following sections:

- 3.2.1.4. Order Items for Search and Data Entry
- 3.2.1.5. Alternative Hierarchical View

3.1.3. A Set of Associations between Components

When SNOMED CT is implemented in electronic health records, there may be situations where explicitly stating associations between components can support effective and efficient use of SNOMED CT.

In some situations, a set of unordered associations between components may be required. In other situations, an ordered list of directed associations between components may be needed. As illustrated in the table below, unordered associations can be represented using an association reference set, while an ordered list of directed associations can be represented using an ordered reference set. Associations can be specified between components of any type. However, associations are typically used to link concepts and/or descriptions.

Table 3.1.3-1: Requirements for a set of associations between components

Requirement	Description	Example Uses	Reference Set
A set of directed associations between components	A set of directed associations between pairs of concepts A set of directed associations between pairs of descriptions	• Grouping concepts together • For example, representing categories of concepts that are used for reporting • Historical associations between components • For example, associating inactive and active duplicate concepts	5.4. Association Reference Set
An ordered list of directed associations between components	An ordered list of directed associations between pairs of concepts or descriptions	Defining alternative hierarchies for navigation and selection of concepts or descriptions. Examples include: • Ordering hierarchical lists of enumerated body structures such as fingers, vertebrae and cranial nerves • Organizing the display of diseases • Commonly seen in a particular specialty.	5.3. Ordered Reference Set

For more detailed use case examples, please refer to the following sections:

- 3.2.1.6. Use Case Specific Associations
- 3.2.6.3.2. Representing Historical Associations
- 3.2.1.5. Alternative Hierarchical View

3.1.4. A Set of Components Annotated with Additional Information

In some cases, an implementer may need to add additional information about each member of a subset. This information may be additional textual information or additional coded values.

Annotating each member of a subset with additional information can facilitate the processing of subset members and assist in meeting the functional requirements of a system.

Table 3.1.4-1: Requirements for a set of components annotated with additional information

Requirement	Description	Example Use	Reference Set
A set of components with free text annotations	A set of concepts, descriptions or relationships each annotated with a free text note	Displaying a textual note for each concept in a list. For example, displaying an advisory note on how to request a particular procedure.	Annotation reference set
A set of components with coded annotations	A set of concepts, descriptions or relationships each annotated with a reference to another component	Marking each concept with a specific coded value to support automated processing of a list. For example, marking each inactive concept with a code that indicates the reason they were inactivated	Attribute value reference set
		Specifying whether descriptions are preferred or acceptable in a given dialect, care setting or clinical context.	Language reference set

For more detailed use case examples, please refer to the following sections:

- Linking concepts to web resources
- Representing reasons for component inactivation
- Indications of acceptability of descriptions

3.1.5. A Set of Maps between SNOMED CT and Another Code System

A map is an association between codes from one code system and codes from another code system, that have the same (or similar) meaning. Mapping is the process of defining a set of maps. Maps are developed in accordance with a documented rationale, for a given purpose. As a result, there may be different maps between the same pair of concepts or terms to meet different use cases.

The purpose of mapping between SNOMED CT and another code system is to provide a link between the code systems, to obtain a number of benefits. These may include:

- Data reuse - for example, SNOMED CT based clinical data can be reused to report statistical and management information using an alternative classification system
- Retaining the value of existing data when migrating to newer database formats and code systems
- Avoiding the need to enter data multiple times and preventing the associated cost and potential errors
- Promoting interoperability between terminologies, classifications and code systems

Table 3.1.5-1: Requirements for a set of maps between SNOMED CT and another code system

Requirement	Description	Example Use	Reference Set
An equivalence map	A set of one-to-one bidirectional maps between SNOMED CT components and codes from another code system	Mapping legacy codes to equivalent SNOMED CT concepts	Simple map reference set
A non-equivalence map	A set of maps from SNOMED CT concepts to codes in another code system, where the map may include: - one-to-many or many-to-one maps - map groups - map rules - map advice	Representing a map from SNOMED CT to a statistical classification	Complex and extended map reference sets
	A set of maps from another code system to SNOMED CT, where the map may include: - one-to-many or many-to-one maps - map groups - map rules - map advice	Representing a map from a statistical classification to SNOMED CT	Complex and extended map reference sets
	A set of links between codes in another code system and SNOMED CT expressions	Representing the link between LOINC codes and SNOMED CT Expressions	Customized reference set

3.1.6. A Set of Sets of Components

Some situations may require sets of components to be implemented and managed. For example, when implementing a package of subsets.

Table 3.1.6-1: Requirements for a set of maps between SNOMED CT and another code system

Requirement	Description	Example Use	Reference Set
A set of sets of components	A set of intensional subset definitions	Managing the set of intensional subset definitions that are used within a particular system, organization, domain or message	Query specification reference set
	A set of references to concepts that represent reference sets	Managing the set of reference sets that are used within a particular system, organization, domain or message	Simple reference set

For more detailed use case examples, please refer to the following sections:

- Managing value sets
- Constraining the coded content of messages

3.2. Use Cases

The table below illustrates some of the use cases that SNOMED CT reference sets support. Furthermore, it includes some typical example uses for each use case, and presents the types of reference sets relevant to these example. You can click on the diagrams and the text to navigate to the subpage, which provides more information about each of the use cases, examples and reference sets. The examples shown in the table are not exhaustive. Instead, they illustrate some common scenarios for each type of reference set.

Table 3.2-1: Reference set use cases

USE CASE	EXAMPLES	REFERENCE SET
Search and Data Entry	• Constrain searches • Constrain data entry • Exclude content	Simple reference set Query specification reference set
	• Order items for search and data entry • Display alternative hierarchy for navigation	Ordered reference set
	• Specify preferred descriptions for display	Language reference set Query specification reference set
Knowledge Linkage	• Linking concepts to web resources • Linking concepts to textual information	Annotation reference set
Reporting and Analytics	• Specifying queries	Simple reference set Query specification reference set
	• Link to statistical classifications	Simple map reference set (specification) Complex and extended map reference set (specification)
Communication	• Constraining the coded content of messages	Simple reference set Query specification reference sets
Language and Dialect	• Indications of acceptability of descriptions	Language reference set Simple reference set
Maintenance and Management	• Consistent replacements of inactive components	Association reference set
	• Managing value sets	Attribute value reference set Association reference set

3.2.1. Search and Data Entry

Many clinical applications include data entry interfaces controlled or assisted by protocols, templates or structured data entry forms. Each field on a data entry form may allow only a limited set of terms or concepts to be entered. The set of candidate term (field) or concept may range from very small (e.g. a set of priorities for a procedure) to very large (e.g. any general diagnosis). Reference sets can be used to restrict the possible values that meet the requirements of a particular data entry protocol.

Examples of using reference sets to support search and data entry include:

- 4.2.1. Simple Reference Set can be used to constrain searches or provide values for selection lists.
- 4.2.2. Ordered Reference Set can be used to prioritize search results or provide an alternative ordering of search results.
- 4.2.6. Annotation Reference Set can be used to supplement search results or data entry options with additional textual or coded information, such as advice on intended usage.
- 5.7. Language Reference Set can be used to ensure that the preferred descriptions, for a given dialect, care setting or clinical context, are displayed.

A SNOMED CT enabled application can use an appropriate reference set to display the valid data entry options and constrain text searches. For more detailed use case examples, please refer to the following sections:

- 3.2.1.1. Constrain Data Entry
- 3.2.1.2. Constrain Searches
- 3.2.1.3. Exclude Content
- 3.2.1.4. Order Items for Search and Data Entry
- 3.2.1.5. Alternative Hierarchical View
- 3.2.1.6. Use Case Specific Associations

3.2.1.1. Constrain Data Entry

In many care settings, similar data sets are collected for each patient. Clinical consultations for many conditions involve repeatable sequences of data entry. These structured and predictable data entry requirements can be met using sets of customized data entry forms designed to collect appropriate data items.

When using a structured data entry mechanism, SNOMED CT encoded data can be selected in a variety of ways. For example, the concepts or descriptions may be selected directly from a list, or the encoding may result from responses to simple choices or the entry of particular values. Simple reference sets can ensure that SNOMED CT codes are entered effectively and consistently.
A simple reference set of concepts may be used to represent the options available in a small selection list. Similarly, a simple reference set of descriptions or a language reference set may specify the set of descriptions available for searching in a specific coded data element. The figure below illustrates how a simple reference set is is used as a value set in a data entry form.
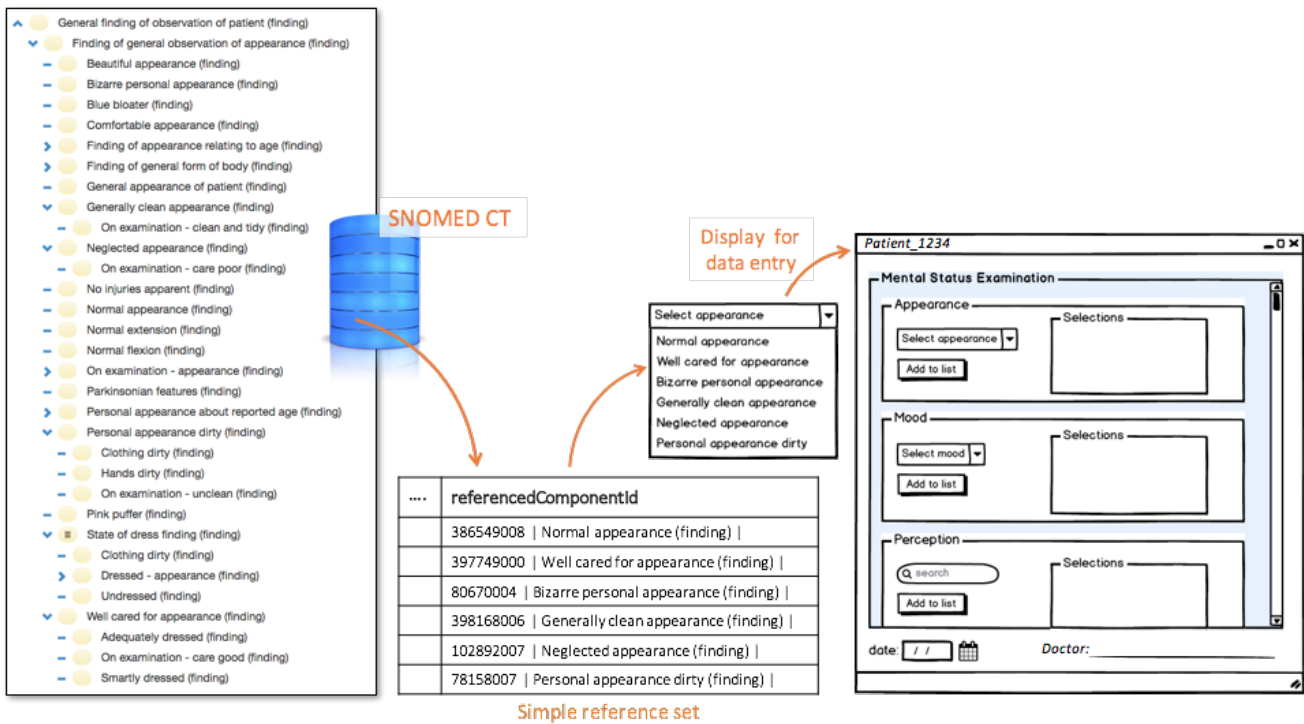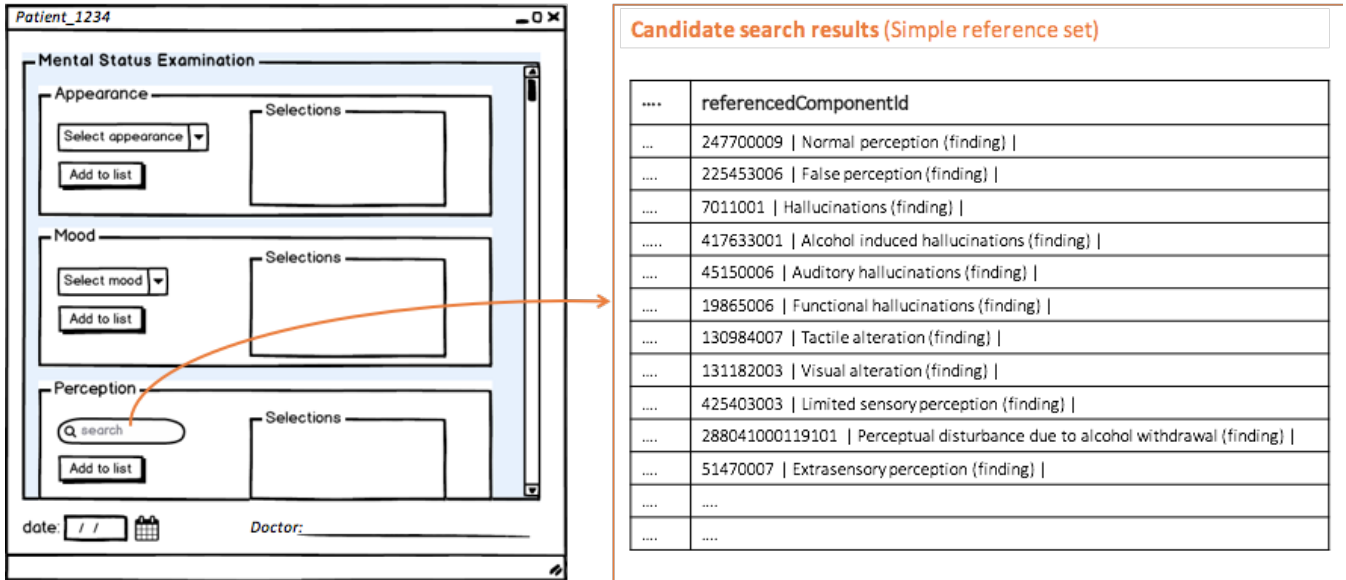
Figure 3.2.1.1-1: Using simple reference sets to constrain data entry

3.2.1.2. Constrain Searches

Simple reference sets enable text searches to be constrained to those components relevant to a particular field. The Search and Data Entry Guide provides additional detail about how to make effective and efficient search capabilities using SNOMED CT. The figure below illustrates the use of a simple reference set to constrain the values returned by a text search in a data entry form. Additionally, dedicated search features support searching the content of the reference set.

Figure 3.2.1.2-1: Using simple reference sets to constrain searches

3.2.1.3. Exclude Content

Even though subsets are typically used to specify content for inclusion, some situations may require particular components to be excluded from another set. Excluding sets of SNOMED CT components can be used to prevent certain concepts appearing in particular search and data entry items.

Like every subset of SNOMED CT components, it is possible to define the subset for exclusion either intensionally or extensionally. This is illustrated in the figure below. When intensionally defined, a query specification reference set can be used for the intensional definition, and a simple reference set can be used for the expansion of the set.
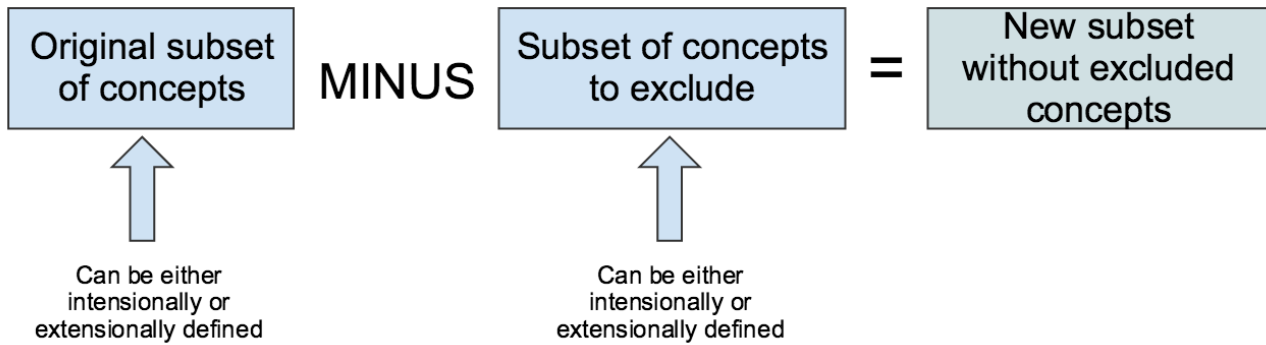
Figure 3.2.1.3-1: Subsets for excluding content can be either intensionally or extensionally defined

---

3.2.1.4. Order Items for Search and Data Entry

The criteria for a successful implementation of SNOMED CT includes the customization of SNOMED CT to meet user needs. The order in which SNOMED CT components are displayed is often important for data entry and searching. This topic is further explored in the Search and Data Entry Guide. In general, rational ordering of selectable items depends on the nature of the application and its operating environment. The table below shows examples of ordering data entry items and search results rationally.

Table 3.2.1.4-1: Examples of rational ordering

Approach	Description	Example Uses	Reference Set
Sequential ordering	Annotating each subset member with an integer, which specify the consecutive order of the members. Two subset members do not have the same number assigned to them.	Displaying descriptions sequentially according to their specified order.	Ordered reference set

Prioritization	Annotating each subset member with a an integer, which specify a priority order. Two or more subset members may have the same number assigned to them.	Showing concepts with a high priority before their siblings using hierarchical display results.
		
		• Display search results in priority order
		  • Results with same rank ordered by shortest or closest match
		• Displaying a rank indicator in search result list
		
		Initially listing concepts and associated descriptions with a priority above a specified threshold and requiring additional steps to access those assigned a lower priority.
		
		• Initial search is conducted on components with highest priority
		• Allow search to be extended to lower priorities
		  • If no high priority matches
		  • If user requests more matches
			Ordered reference set

Sequential Ordering

Displaying items for data entry in a rational way typically involves organizing the values in a selection list in an order that is logical for the end users. As illustrated in the figure below, an ordered reference set can be used to specify the order in which SNOMED CT components should be displayed.

Figure 3.2.1.4-1: Example of how an ordered reference set can be used to order items in a drop down list

Examples of presenting concepts (or descriptions) in an order that is rational or helpful for a particular purpose include:

- Displaying numbered body parts, such as fingers, cranial nerves or vertebrae, in numeric order
- Displaying ordinal values, such as frequencies, severities or stages, from lowest to highest

The table below shows how the order of cranial nerves can be specified in an ordered reference set. The order attribute is used to indicate the sequential order of each subset member. Note that the linkedToId is set to 0 in this case because this reference set does not specify a hierarchy of reference set members.

Table 3.2.1.4-2: An ordered reference set used to specify the order of cranial nerves

refsetId	referencedComponentId	order	linkedToId
609999999102 |Cranial nerve simple reference set|	11522000 |Olfactory nerve structure (body structure)|	1	0
609999999102 |Cranial nerve simple reference set|	18234004 |Optic nerve structure (body structure)|	2	0
609999999102 |Cranial nerve simple reference set|	56193007 |Oculomotor nerve structure (body structure)|	3	0
609999999102 |Cranial nerve simple reference set|	39322007 |Trochlear nerve structure (body structure)|	4	0
609999999102 |Cranial nerve simple reference set|	80622005 |Abducens nerve structure (body structure)|	5	0
609999999102 |Cranial nerve simple reference set|	27612005 |Trigeminal nerve structure (body structure)|	6	0
609999999102 |Cranial nerve simple reference set|	56052001 |Facial nerve structure (body structure)|	7	0
609999999102 |Cranial nerve simple reference set|	8598002 |Vestibulocochlear nerve structure (body structure)|	8	0
609999999102 |Cranial nerve simple reference set|	21161002 |Glossopharyngeal nerve structure (body structure)|	9	0
609999999102 |Cranial nerve simple reference set|	88882009 |Vagus nerve structure (body structure)|	10	0

609999999102 |Cranial nerve simple reference set|	15119000 |Accessory nerve structure (body structure)|	11	0
609999999102 |Cranial nerve simple reference set|	37899009 |Hypoglossal nerve structure (body structure)|	12	0

If there is a need to specify a customized hierarchical structure to support navigation, this can be achieved by specifying an alternative hierarchical view using an ordered reference set.

Prioritization

Some situations may require a set of subset members to be grouped. For example, a set of concepts may need to be grouped based on how frequently they are used within a particular specialty, department or data entry scenario. In this case, an ordered reference set may be used for prioritization, instead of a purely sequential ordering of each member. Prioritization is similar to sequential ordering, but also supports assigning the same rank to multiple components. A common use of prioritization is to support rational ordering of concepts or descriptions for display of data entry items and search results. More advanced uses may also be required, for example where the priority order is used to trigger certain decision support features or data entry options.

referencedComponentId	order
1225002	radiography of humerus		1
1597004	skeletal X-ray of ankle and foot		2
168594001	clavicle X-ray		2
168619004	plain X-ray head of humerus		2
168620005	plain X-ray shaft of humerus		2
168623007	X-ray shaft of radius/ulna		2
168637003	plain X-ray radius		2
168655007	instability views carpus		2
168663008	plain X-ray head of femur		2
168664002	femoral neck X-ray		2
168665001	plain X-ray shaft of femur		2
168669007	patella X-ray		2
241063007	bicipital groove X-ray		2
241066004	ulna groove X-ray		2
241069006	ulna X-ray		1
241071006	scaphoid X-ray		1
241073009	metacarpal X-ray		1
241075002	femur X-ray		1
241076001	tibia and/or fibula X-ray		2
241077005	tibia X-ray		1
241078000	fibula X-ray		1
241079008	metatarsal X-ray		1
241080006	tarsus X-ray		1
268427003	X-ray shaft of tibia/fibula		2
271311001	carpal bones X-ray		2
302402006	radius and/or ulna X-ray		2
37815002	diagnostic radiography of calcaneus		2
40348008	skeletal X-ray of pelvis and hip		2
418687005	fluoroscopy of humerus		2
427961005	x-ray of acetabulum		2
205115004	radiologic examination of femur, anteroposterior and lateral views		2
432552002	computed tomography of clavicle		2
48966008	skeletal X-ray of shoulder and upper limb		2
5433008	skeletal X-ray of lower limb		1
70780000	skeletal X-ray of elbow and forearm		2
72872009	skeletal X-ray of upper limb		1
79082005	diagnostic radiography of fibula, combined AP and lateral		2
82420003	radiologic examination of forearm, anteroposterior and lateral views		2

1) Display most frequently used options first →

1225002	radiography of humerus	
241069006	ulna X-ray	
241071006	scaphoid X-ray	
241073009	metacarpal X-ray	
241075002	femur X-ray	
241077005	tibia X-ray	
241078000	fibula X-ray	
241079008	metatarsal X-ray	
241080006	tarsus X-ray	
5433008	skeletal X-ray of lower limb	
72872009	skeletal X-ray of upper limb	

2) Expand list to show all options →

1597004	skeletal X-ray of ankle and foot	
168594001	clavicle X-ray	
168619004	plain X-ray head of humerus	
168620005	plain X-ray shaft of humerus	
168623007	X-ray shaft of radius/ulna	
168637003	plain X-ray radius	
168655007	instability views carpus	
168663008	plain X-ray head of femur	
168664002	femoral neck X-ray	
168665001	plain X-ray shaft of femur	
168669007	patella X-ray	
241063007	bicipital groove X-ray	
241066004	ulna groove X-ray	
241076001	tibia and/or fibula X-ray	
268427003	X-ray shaft of tibia/fibula	
271311001	carpal bones X-ray	
302402006	radius and/or ulna X-ray	
37815002	diagnostic radiography of calcaneus	
40348008	skeletal X-ray of pelvis and hip	
418687005	fluoroscopy of humerus	
427961005	x-ray of acetabulum	
205115004	radiologic examination of femur, anteroposterior and lateral views	
432552002	computed tomography of clavicle	
48966008	skeletal X-ray of shoulder and upper limb	
70780000	skeletal X-ray of elbow and forearm	
79082005	diagnostic radiography of fibula, combined AP and lateral	
82420003	radiologic examination of forearm, anteroposterior and lateral views	

Figure 3.2.1.4-2: Using a priority order to display data entry options

ⅰ The refsetId used in this table is fictitious

3.2.1.5. Alternative Hierarchical View

Ordered reference sets can be used to specify and display a customized navigation hierarchy. Alternative hierarchical representations of SNOMED CT can support data entry by satisfying the requirements of a specific use case, and addressing some of the challenges of displaying an unordered polyhierarchy (as defined by SNOMED CT's subtype structure).

The figure below shows the way a navigation hierarchy is represented. The example reference set contains a set of description components used to describe finger structures.

The | All fingers | components is linked to the | Hand |, and the | Thumb | is linked to the | All fingers component | The | Thumb | is placed first because it has the order value 1. Similarly, the components for | Second finger |, | Third finger |, | Fourth finger | and | Fifth finger | are also linked to the | All finger | component in the order specified by the order value. As shown in the figure the direction of the associations goes from the referenceComponentId to the linkedToId, so the components referenced by the linkedToId are used to form the groups specified in the hierarchy

id	effective Time	active	moduleId	refsetId	refsetId_term	referencedComponentId	referencedComponentId_term	order	linkedToId	linkedT
...	20160731	1	19999999103	159999999105	Associations as ordered reference set	70327001	All fingers	1	141819019	Hand
...	20160731	1	19999999103	159999999105	Associations as ordered reference set	127053016	Thumb	1	70327001	All finge
...	20160731	1	19999999103	159999999105	Associations as ordered reference set	138873019	Second finger	2	70327001	All finge
...	20160731	1	19999999103	159999999105	Associations as ordered reference set	108884010	Third finger	3	70327001	All finge
...	20160731	1	19999999103	159999999105	Associations as ordered reference set	136021011	Fourth finger	4	70327001	All finge
...	20160731	1	19999999103	159999999105	Associations as ordered reference set	21356012	Fifth finger	5	70327001	All finge

Figure 3.2.1.5-1: Navigation hierarchy example.

The usability of the ordered reference set for representing alternative hierarchy can be maximized by:

- Constraining the number of levels in the hierarchy and/or the number of concepts at each level.
  - Using many levels, each with a relatively small number of concepts, allows the most common options to be displayed with a higher priority.
  - Using fewer levels, each with a relatively large number of concepts can reduce the number of levels that needs to be navigated to find an appropriate concept.
  - Options that are never (or rarely) used can be excluded from a customized navigation hierarchy to limit the range of choices available.

- Ordering each concept at the same hierarchical level, to match user preferences or to facilitate faster access to more frequently used options.
- Ensuring that the navigation hierarchy is adapted to meet the requirements of a specific use case, without affecting the correctness of the subtype hierarchy (and associated logical inferences).

3.2.1.6. Use Case Specific Associations

SNOMED CT represents relationships between concepts that are necessarily (i.e. always) true. However, other relationships between concepts may exist in specific situations or use cases. An 5.4. Association Reference Set can be used to represent these additional relationships, which are not necessarily true, but which are needed for a specific purpose. Examples include:

- Associations between procedures and the clinical findings that serve as an indication for that procedure. These associations enable relevant procedures to be displayed when specific clinical findings are selected.
- Associations between a medication and its known side effects. These associations enable relevant side effects to be displayed when specific medications are selected
- Associations between a disease and the set of possible symptoms that may be experienced. These associations enable relevant diseases to be displayed when a set of symptoms are selected.

5.4. Association Reference Set can be used to constrain (or guide) data entry into fields, where the value is dependent on (or has some type of association with) the value of another field. While other technical solutions are possible, the 4.2.5. Association Reference Set provi des a standardized way of representing and distributing the associations required to support this functionality. The figure below illustrates how an 4.2.5. Association Reference Set could be used for this purpose.
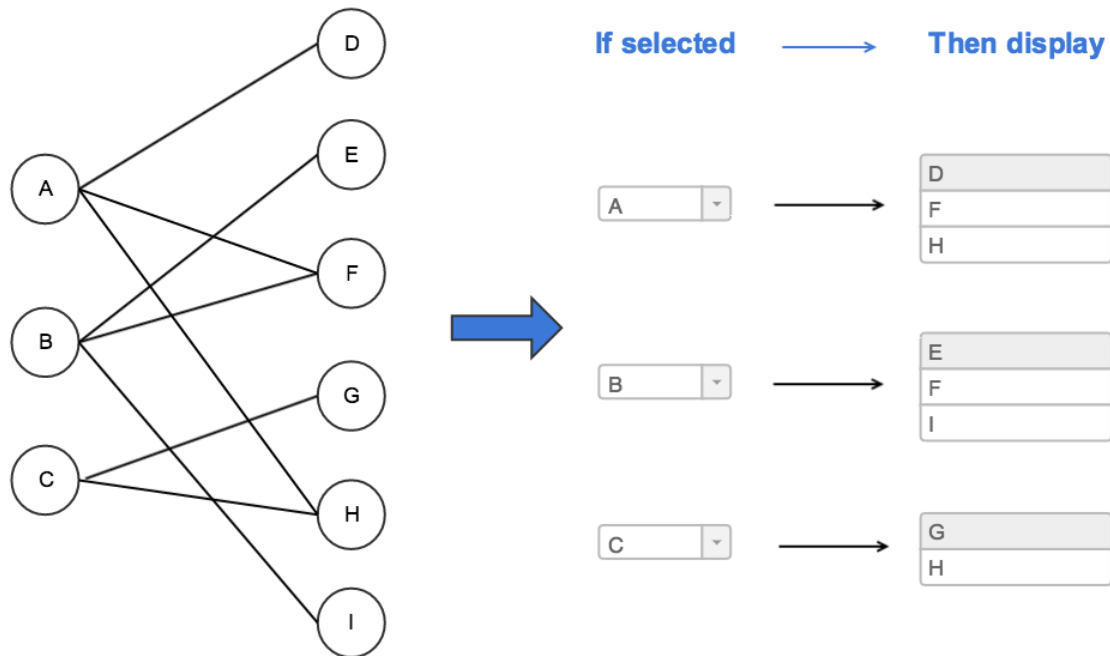
Figure 3.2.1.6-1: Using associations to define dependencies between fields

3.2.2. Knowledge Linkage

Linking SNOMED CT components to knowledge resources (such as clinical guidelines or decision support systems) is a way of adding significant value to electronic health records.

The diagram below illustrates the range of use cases for knowledge linkage including presenting alerts to the user, displaying relevant clinical guidelines and treatment protocols, or automatically populating an order, message or report.
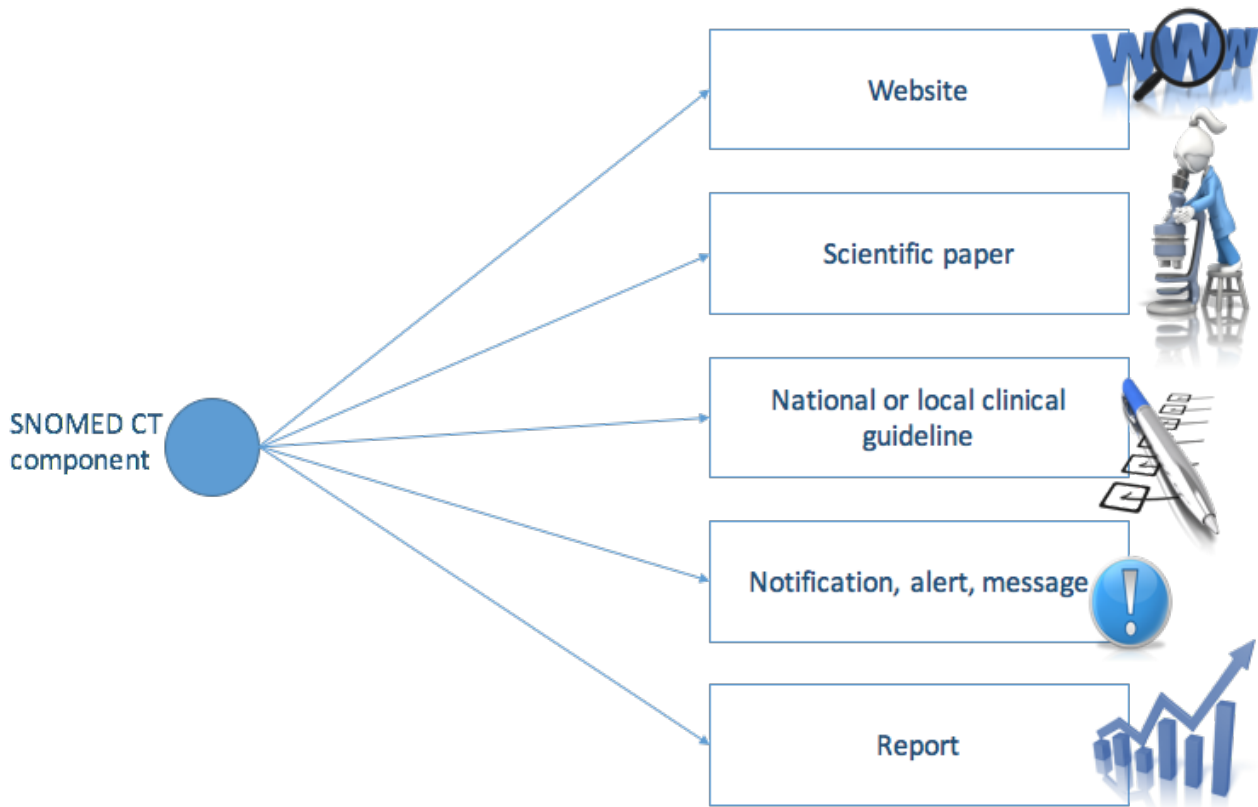
Figure 3.2.2-1: Components annotated with additional information can support different levels of knowledge linkage

Reference sets can be used to enable knowledge linkage in different ways. Examples include but are not limited to:

- Annotation reference sets can be used to add a linkage between a referenced component and a string-representation of a specific knowledge resources.
- Association reference sets can be used to create associations between SNOMED CT components to enable documentation and/or decision support, using the associations as a simple rule-representation.
- Simple map reference sets, complex and extended map reference sets can be used to define maps from other code systems to SNOMED CT, and function as a linkage to sources of non SNOMED CT-encoded information

For a more detailed use case example, please refer to the following section:

- 3.2.2.1. Linking Concepts to Web Resources

3.2.2.1. Linking Concepts to Web Resources

In the example below URLs are used to annotate two SNOMED CT concepts with images on the web. It is not recommended to use this approach to annotate concepts with text that may require translation to other languages. Instead, such text should be included under an appropriate description type within the description file. However, in some cases this type of reference set can be useful to link concepts or descriptions to specific guidance documents or other types of knowledge resources.

Table 3.2.2.1-1: Example of an associated image annotation reference set

refsetId	referencedComponentId	Annotation
900000000000517004 |Associated image|	80891009 |Heart structure|	http://en.wikipedia.org/wiki/Heart#mediaviewer/-File:Wiki_Heart_Antomy_Ties_van
900000000000517004 |Associated image|	86174004 |Laparoscope|	http://www.mdguidelines.com/images/Illustrations/laparosc.jpg

3.2.3. Reporting and Analytics

The main benefits of using an EHR accrue with the implementation of effective retrieval, analysis and reuse of clinical information. Analysis of health record data may cover:

- Individual patient records to search for significant patterns that may prompt interventions
- Patient groups or cohorts, based on demographics, diagnoses, treatments or interventions
- Enterprise groups, based on teams, wards, clinics, institutions or providers
- Geographical groups, based on a local area, town, region or country

SNOMED CT has a number of unique features, which makes it capable of supporting a range of retrieval and analytics functions, which use reference sets. Examples include, but are not limited to:

- Simple reference sets can be used to represent subsets of SNOMED CT concepts which can be used in queries to identify clinical records
- Simple reference sets can be used to represent non-standard aggregations of concepts for specific use cases
- Simple map reference sets, complex and extended map reference sets can be used to define maps from other code systems to SNOMED CT so that clinical data can be prepared for analytics, and then performed using SNOMED CT
- Simple reference sets and ordered reference sets can be used to define language or dialect specific sets of descriptions over which lexical searches can be performed

The document Data Analytics with SNOMED CT provides detailed information about how SNOMED CT can be used for analytics.

For more detailed use case examples, please refer to the following sections:

- 3.2.3.1. Maps to Statistical Classifications
- 3.2.3.2. Specifying Queries for Retrieval and Analysis

3.2.3.1. Maps to Statistical Classifications

Clinical information recorded using SNOMED CT may include data that is relevant to reports, statistical returns, billing claims, etc. that need to be encoded using a specific code system or a statistical classification such as ICD-10. Mapping allows relevant information to be used for those purposes, minimizing the requirement for additional manual data entry.

Maps are represented as reference sets, which are either of type simple, complex and extended map reference sets. Special cases may also occur, which require a customized reference set to represent the map. An example of this is the |LOINC Term to Expression reference set| , which is used to link LOINC Terms to SNOMED CT expressions. The standard reference set format do not support maps to SNOMED CT expressions. Hence, the type of map reference set to use depends on the features that need to be supported by the map.

Simple map reference sets support mapping SNOMED CT codes to a single code or a combination of codes in a target code system. However simple maps are usually only appropriate where there is an equivalent map between SNOMED CT and the values in the other code system.

Complex and extended map reference sets enable the representation of:

- Maps from a single SNOMED CT concept to a combination of codes (rather than a single code) in the target scheme
- Maps from a single SNOMED CT concept to choice of codes in the target scheme. In this case, the resolution of the choices may involve:
  - Manual selection supported by advisory notes
  - Automated selection based on rules that test other relevant characteristics in the source data (e.g. age and sex of the subject, presence or absence of co-existing conditions, etc.)
  - A combination of automated processing with manual confirmation or selection where rules are insufficient to make the necessary decisions

The completeness of mapping between two code systems depends on the scope, level of detail provided by the two code systems and the precision of mapping required to safely meet the intended mapping use case.

The figure below shows an except of some of the reference sets for maps between SNOMED CT and other code systems, which are available with the International Edition of SNOMED CT. However, local maps may also be developed and applied as part of an Extension to SNOMED CT.
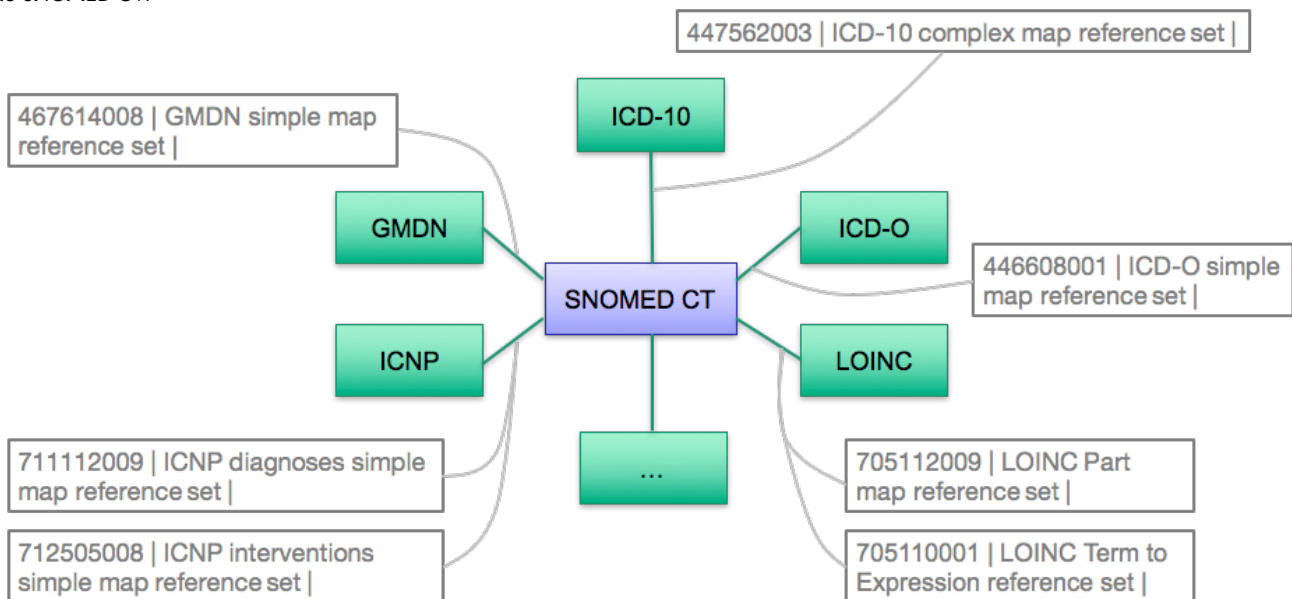
Figure 3.2.3.1-1: Excerpt of mappings in the SNOMED CT International Edition

3.2.3.2. Specifying Queries for Retrieval and Analysis

Both extensionally and intensionally defined subsets of SNOMED CT components are useful for specifying clinical queries. For example, subsets of SNOMED CT concepts can be used to categorize patient data by testing for membership in a predefined subset, which is represented as a reference set.

The SNOMED CT Expression Constraint Language (ECL) enables simple queries over SNOMED CT content to be expressed. While the language itself does not support querying over the full EHR content, the ECL could be embedded within record-based query languages (such as SQL) to represent the terminological aspects of these queries.

Related to reference sets, the ECL includes the ability to refer to a set of concepts that are referenced by members of a reference set. Additionally, it includes a range of features, such as refinements, disjunction, and conjunction, which support specialized queries. The me mberOf function evaluates to the set of concepts that are referenced by the given reference set. For example, the following expression constraint is satisfied by the set of concepts which are members of |Example problem list simple reference set| :

- memberOf 649999999104 |Example problem list simple reference set|

The diagram below illustrate how reference sets can be for specifying queries.
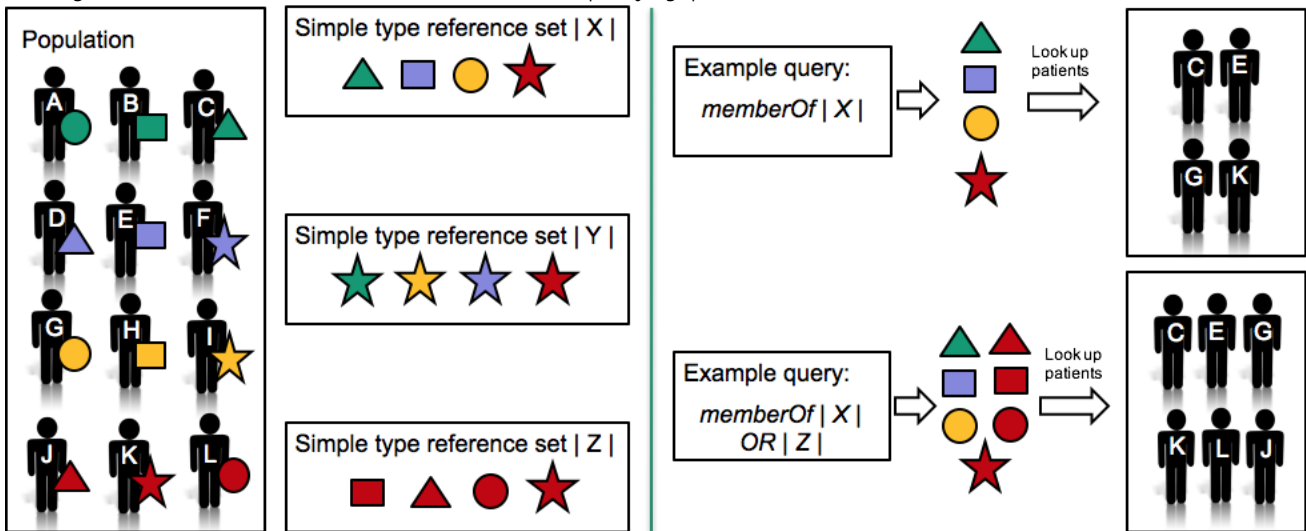
Figure 3.2.3.2-1: Using reference sets for specifying queries

3.2.3.2.1. Categorizing Patients Using Subsets

Subsets of SNOMED CT concepts can be used to categorize patient data by testing for membership in a predefined subset. The diagram below illustrates the use of a simple reference set which includes references to rare disease concepts. SNOMED CT does not currently have a defined mechanism to distinguish rare diseases, so this simple reference set is defined extensionally (i.e. by enumeration). This simple reference set is used to create a cohort, by categorizing patients according to their SNOMED CT encoded records, and specific values of contextual metadata. By comparing the patient diagnosis with the concepts included in the reference set it is possible to count the total number of patients who suffer from rare diseases.

Example Patient Data

Gender:	248152002	Female (finding)	
Date of birth:	19900326
Diagnosis:	82725007	Progressive myositis ossificans	

Gender:	248153007	Male (finding)	
Date of birth:	19540406
Diagnosis:	86406008	Human immunodeficiency virus infection	

Gender:	248152002	Female (finding)	
Date of birth:	19660817
Diagnosis:	19138001	Epidermodysplasia verruciformis	

Gender:	248152002	Female (finding)	
Date of birth:	20010213
Diagnosis:	193462001	Insomnia (disorder)	

Gender:	248153007	Male (finding)	
Date of birth:	19610102
Diagnosis:	86514004	Granulomatous hepatitis (disorder)	

Gender:	248153007	Male (finding)	
Date of birth:	20090730
Diagnosis:	238870004	Hutchinson-Gilford syndrome	

Gender:	248152002	Female (finding)	
Date of birth:	19771023
Diagnosis:	50440006	Crohn's disease of colon	

Patient data storage

Does the patient have a diagnosis which is included in the 'Rare diseases simple reference set'?

Yes — Count: 3

No — Count: 4

....	referencedComponentId
...	82725007	Progressive myositis ossificans (disorder)	
....	357705009	Cotard's syndrome (disorder)	
....	74774004	Urticaria caused by cold (disorder)	
.....	19138001	Epidermodysplasia verruciformis (disorder)	
....	44600005	Xeroderma pigmentosum (disorder)	
....	240820001	Lymphatic filariasis (disorder)	
....	29966009	Hypertrichosis (disorder)	
....	89870006	Aquagenic urticaria (disorder)	
....	238870004	Hutchinson-Gilford syndrome (disorder)	
....	62732009	Alien limb phenomenon (disorder)	
....	....
....	....

Example of a Rare Diseases Simple Reference Set

Figure 3.2.3.2.1-1: Categorize patients using a simple reference set

3.2.3.2.2. Categorization of Patient Data Using Subsumption Testing

A combination of reference sets and subsumption testing can be used to enable a simple, yet sophisticated, analytics feature. For example, you may want to include all of the descendants of reference set members in a particular analysis.

The diagram below illustrate the difference of categorizing patients using two approaches:

1. A subset of components.
    - Two patients are included in a query which analyzes encoded health records and checks for membership in the appropriate subset. E.g. patients with a finding included in the rare diseases reference set.
2. A subset of components and subsumption testing.
    - Three patients are included in a query which test the codes recorded in patient records and check for membership in the appropriate subset or descendants of the subset members. E.g. patients with an associated finding that is referenced in the rare diseases reference set, or any subtypes of the concepts in the reference set.
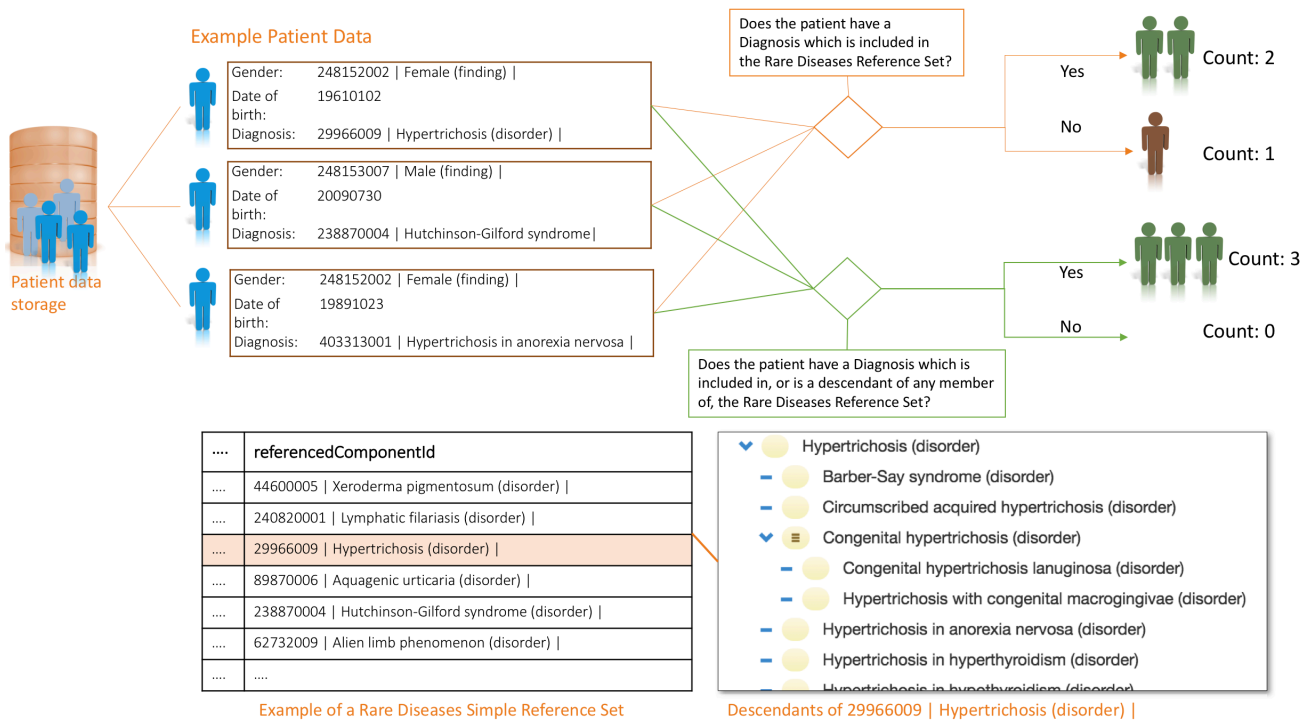
Figure 3.2.3.2.2-1: Categorizing patients using a simple reference set and subsumption testing

3.2.4. Communication

Messages and communication services are a means of exchanging data and thus enable effective and efficient communication among healthcare professionals and between patients and providers. SNOMED CT is important for communication because it serves as a semantic foundation for the meaning expressed in a message. Hence, SNOMED CT can ensure consistent and accurate representation of the information communicated, and support correct interpretation of the clinical information within a message.

Communicating clinical data through messages support a range of purposes, including:

- Delivering accurate, accessible, and actionable health information that is targeted or tailored.
- Facilitating the meaningful use and exchange of health information among healthcare professionals.
- Supporting shared decision-making between patients and providers.
- Providing personalized self-management tools and resources.
- Building social support networks.
- Increasing health literacy skills.

Use of SNOMED CT in Messages

Healthcare messages include fields that can be populated with codes from clinical coding schemes. SNOMED CT provides concept identifiers as a means of encoding concepts . These concept identifiers are suitable for use in appropriate fields of many clinical messages.

Implementations of clinical messaging typically constrain the range of values that can be applied to particular fields several reasons for this are listed in the following table.

Table 3.2.4-1: Reasons for constraining the content of fields in clinical messages

Reason	Example
To ensure that the information encoded is meaningful as a value for the specified field.	A field that is intended to describe the nature of investigation may contain a code that means "Serum glucose measurement" but should not contain a code that means "Hypoglycemia."
To ensure that receiving application is able to process the message.	A locally added code value may be valid in a particular application but should not be used if the receiving application needs to retrieve, process or analyze the coded part of the message.
To ensure adequate detail and specificity.	A field used to report an operative procedure could contain a code for "Abdominal procedure." However, this would not be adequate to meet the business purpose served by a message.
To avoid unnecessary detail or diversity.	A biochemical investigation could be reported using a code that represents various detailed aspects of the method used to perform the investigation. Such details may be unnecessary to a clinician and may complicate the analysis, charting and graphing of a series of results reported at different levels of detail.

For a more detailed use case example, please refer to the following section:

- 3.2.4.1. Constraining the Coded Content of Messages

3.2.4.1. Constraining the Coded Content of Messages

Communication specifications define structures designed to meet particular requirements . For example, recording a decision to prescribe a particular pharmaceutical product or substance might trigger an electronic prescription sent to the pharmacy. Reference sets may also be used to specify the allowed values in messages and for constraining the codable elements in data entry models. For some bindings it may be relevant to apply certain conditions, to enable that one value set is displayed given a specified criteria, and another value set is displayed given another criteria (or set of criterions). An example of such conditional value set binding [1] is illustrated below.
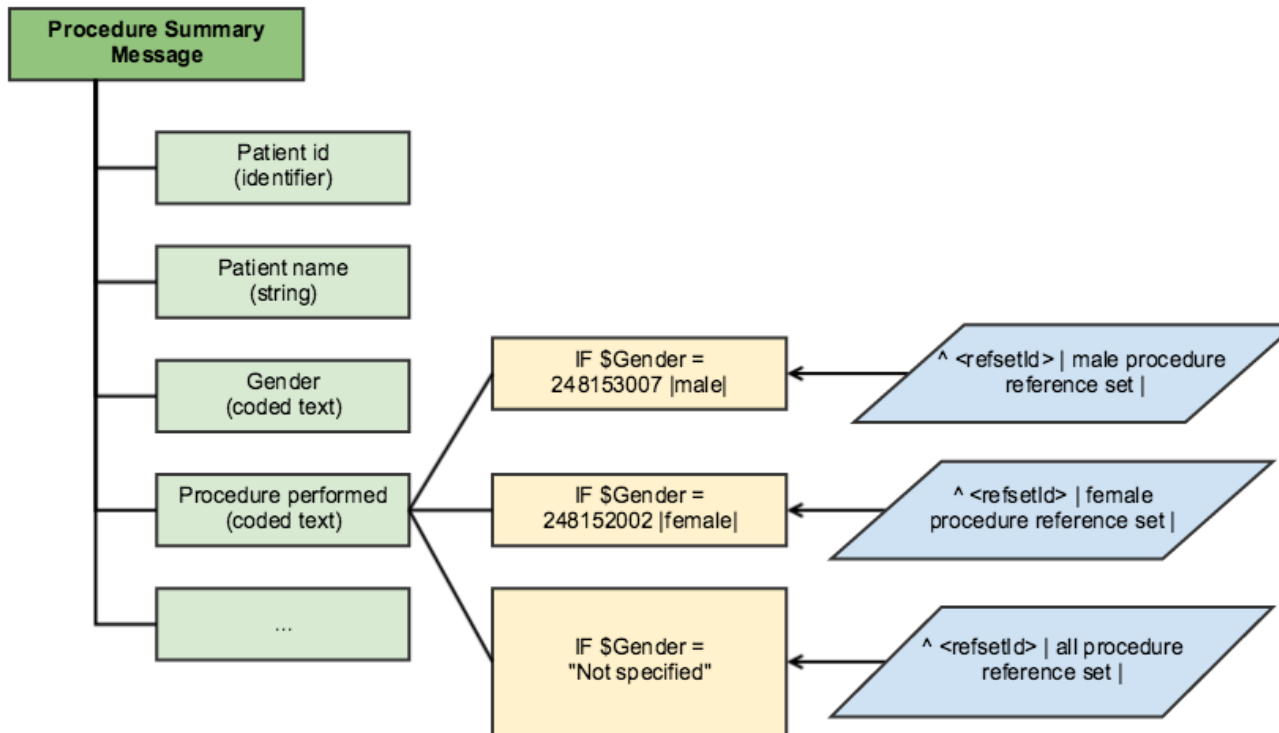
Figure 3.2.4.1-1: Conditional value set binding

Table 3.2.4.1-1: Types of Reference Sets applicable for messages

Type	Description
Simple reference sets	A simple reference set may be used to represent a SNOMED CT-based value set applicable to a particular field in a message. The items to be populated in a particular field in the message can be constrained by filtering searches so that only concepts within that reference set are returned.
Query specification reference set	Query specification reference sets may be used to represent a set of intensionally defined SNOMED CT subsets, where each subset represents the value set for a particular field in a communication messages. One query specification reference set may therefore be used to hold all value sets applicable within a single communication messages, or within a set of messages.

[1] Value set bindings are used to express the valid values used to populate an information model artifact. For example the value set used to populate a drop down menu in a user interface or the valid values for a coded text in a message model.

3.2.5. Language and Dialect

SNOMED CT is a multinational, multilingual terminology which enables the link between concepts and different linguistic representations. Hence, SNOMED CT has a built-in framework to manage different languages and dialects. Each concept, representing a clinical meaning, can be linked to descriptions, which express that particular meaning using different terms, languages or dialects. This means that descriptions can be added to SNOMED CT concepts and expressed in other languages than what is included in the International Edition. Often these additional descriptions are used in a national Extension of SNOMED CT.

Consequently, one aspect of implementing and customizing SNOMED CT to meet specific user needs is to determine and specify preferences for the descriptions to be used, e.g. for display in an electronic health record or more generally, as preferences within a specific clinical domain or facility.

In the International Edition of SNOMED CT language reference sets are included which specifies the preferred and acceptable synonyms for each concept in both US english and GB english. However, language reference sets can also be developed and applied locally to specify what descriptions are preferred and acceptable in a given context. This means, that even within a single country, or a single hospital, different descriptions can be applied to meet the user preferences, even without relaxing the need for consistency and unambiguous concept definitions.

For more detailed use case examples, please refer to the following sections:

- 3.2.5.1. Indications of Acceptability of Descriptions
- 3.2.5.2. Interface Terminology

3.2.5.1. Indications of Acceptability of Descriptions

When SNOMED CT is being used as an interface terminology, the preferred term for each concept should be used as the default for display on the user interface. Each concept may have a different preferred term in different languages, dialects, specialties or care settings, and so these can be configured for a specific clinical environment. Preferred terms and acceptable synonyms are defined in SNOMED CT using a language reference set, which references the subset of descriptions used in a given language, dialect, specialty or care setting. Two language references sets are distributed with the International Edition of SNOMED CT (for US-English and UK-English), and various member countries distribute their own national language reference sets. Additional language reference sets may be created at the regional, specialty, institute or software product level to truly customize the local user's experience.

The most common use case for language reference sets is to specify the acceptable and preferred terms for use within a particular country or region. As illustrated below, descriptions associated with a single SNOMED CT concept can be specified as a preferred or acceptable synonym.
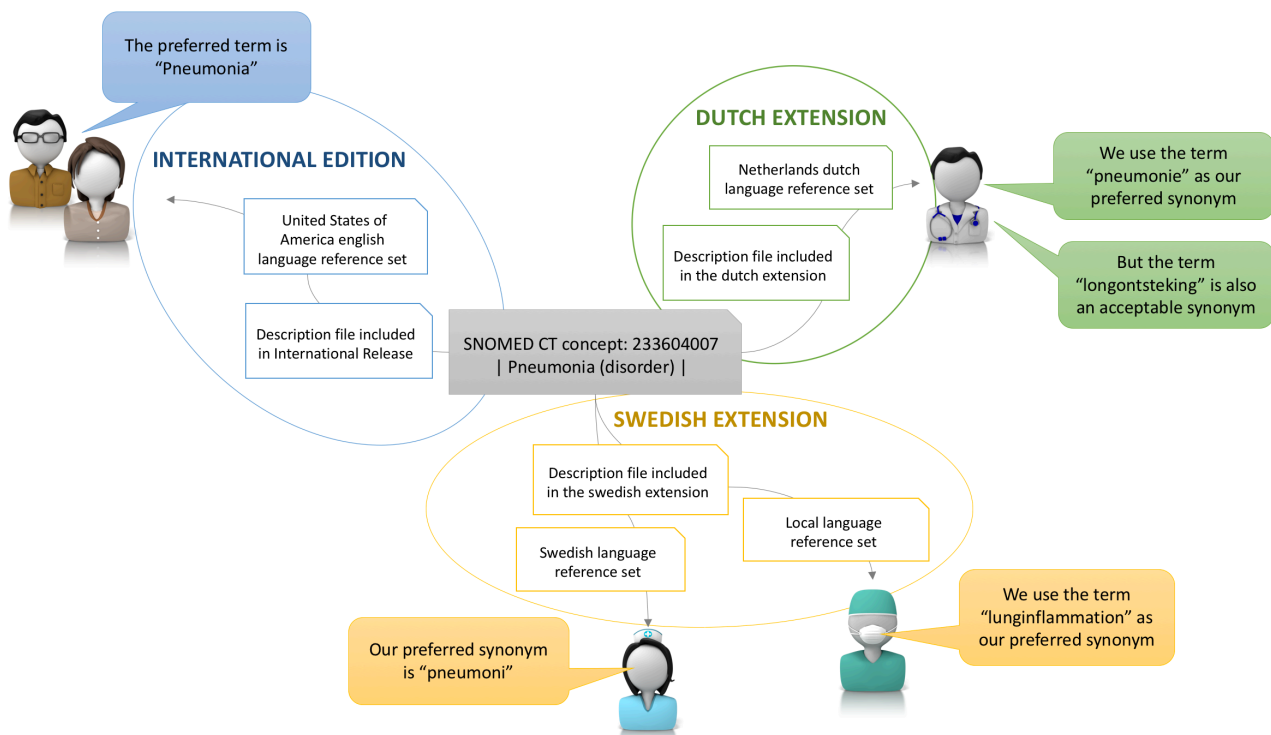
Figure 3.2.5.1-1: Language reference sets and its relation to Description files

3.2.5.2. Interface Terminology

Using SNOMED CT at the user interface level can be accomplished in various ways using different types of reference sets. Approaches include:

- Using SNOMED CT directly as interface terminology
    - using a simple reference set of descriptions used for display
    - using a language reference set which specifies preferred and acceptable synonyms
- Using a separate interface terminology which is mapped to SNOMED CT
    - using a simple map reference set to represent the linkage between the interface terminology and SNOMED CT concepts.

Using SNOMED CT as an Interface Terminology

Simple Reference Set of Descriptions

As SNOMED CT descriptions can be used directly as an interface terminology, subsets of SNOMED CT descriptions may be directly shown to a clinician at the user interface. These description subsets may be customized

- for a specific language or dialect (such as Spanish or Australian English)
- for a given clinical specialty (such as cardiology or oncology)
- for a user type (such as a doctor, nurse or patient)
- a care setting (such as an aged care home or a hospital inpatient ward) or
- a specific document or field in a health record

For each of these use cases, the set of acceptable terms and the preferred term for that clinical use can be identified. However, when searching SNOMED CT it is recommended that any term that is considered "acceptable" for use in a given context should be available to support searching for an appropriate concept, while the preferred term is often used to confirm the intended meaning of the selection.

Representing a subset of descriptions can be done using a simple reference set, as illustrated in the diagram below. However, other types of reference set types may be feasible if additional features are required, such as specifying which descriptions are preferred or acceptable (language reference set), ordering or prioritizing the descriptions (ordered reference set), annotating the description with some textual information (annotation reference set) or associating the descriptions to other components (association reference set).
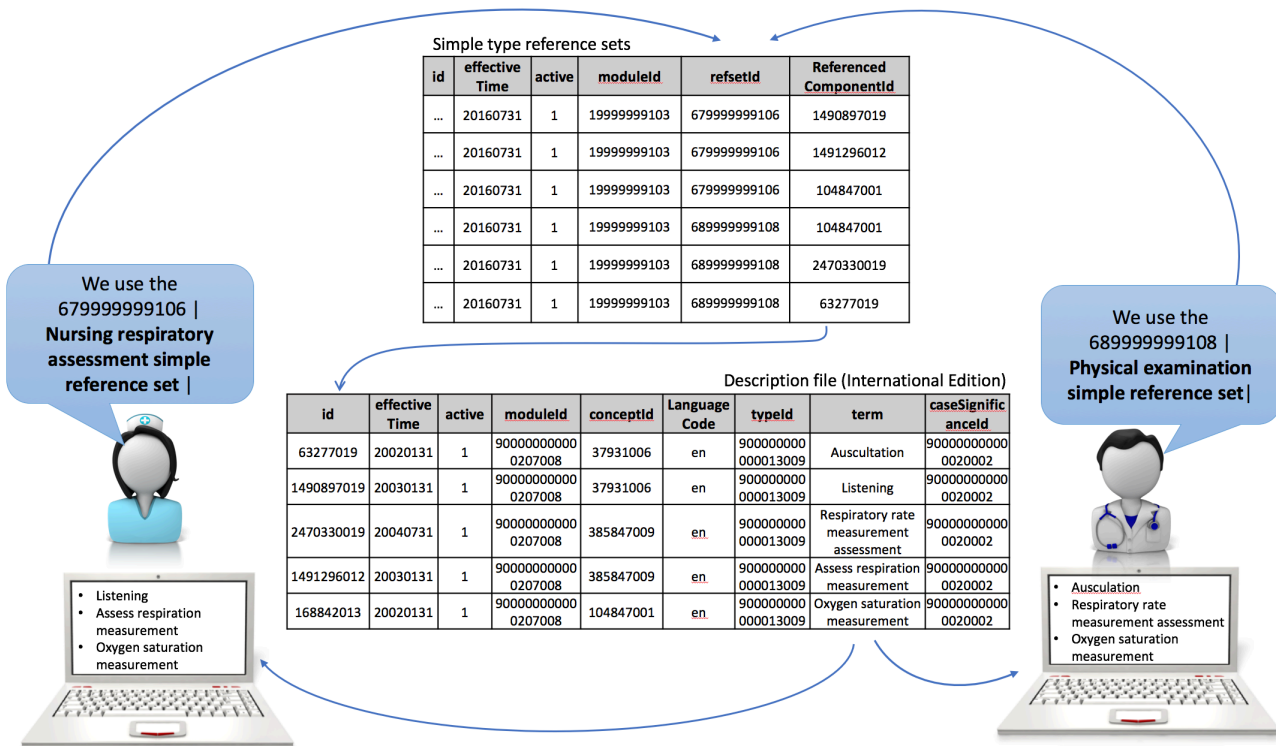
Figure 3.2.5.2-1: Using simple reference set of descriptions to specify terms for display.

Language Reference Set

While a simple type reference set can be used for defining a subset of descriptions, a language reference set is designed to support indication of language and dialect preferences through the addition of the 'acceptability' attribute. This allows preferred and acceptable

descriptions to be defined for any context of use, including within a particular country or region, within a clinical specialty or care setting, within an organization or department, or for a specific type of user.

The table below shows an excerpt from two language reference sets, which are both distributed with the International release of SNOMED CT, i.e. the 900000000000509007 |United States of America English language reference set| and the 900000000000508004 |Great Britain English language reference set| . Both reference sets reference Descriptions available in the Description file in the International Edition.

Table 3.2.5.2-1: Excerpt from the | United States of America English language reference set | and the | Great Britain English language reference set |, which are both distributed with the International release of SNOMED CT.

id	effective Time	active	moduleId	moduleId_term	refsetId	refsetId_term	referencedComponentId	ReferencedComponentId_t
...	20160731	1	900000000000207008	SNOMED CT core module	900000000000509007	United States of America English language reference set	132967011	Appendectomy
...	20160731	1	900000000000207008	SNOMED CT core module	900000000000509007	United States of America English language reference set	132972019	Excision of appendix
...	20160731	1	900000000000207008	SNOMED CT core module	900000000000508004	Great Britain English language reference set	132972019	Excision of appendix
...	20160731	1	900000000000207008	SNOMED CT core module	900000000000508004	Great Britain English language reference set	132973012	Appendicectomy

The diagram below illustrate the use of language reference sets and show the relation between the language reference set and the description file. Even though three descriptions are specified for the same concept in the description file, the language reference set specifies which of these descriptions are preferred and acceptable within a given language, dialect or organisation. If a description is not referenced in the language reference set, then that particular description can be regarded as not acceptable within the context where the language reference set apply.
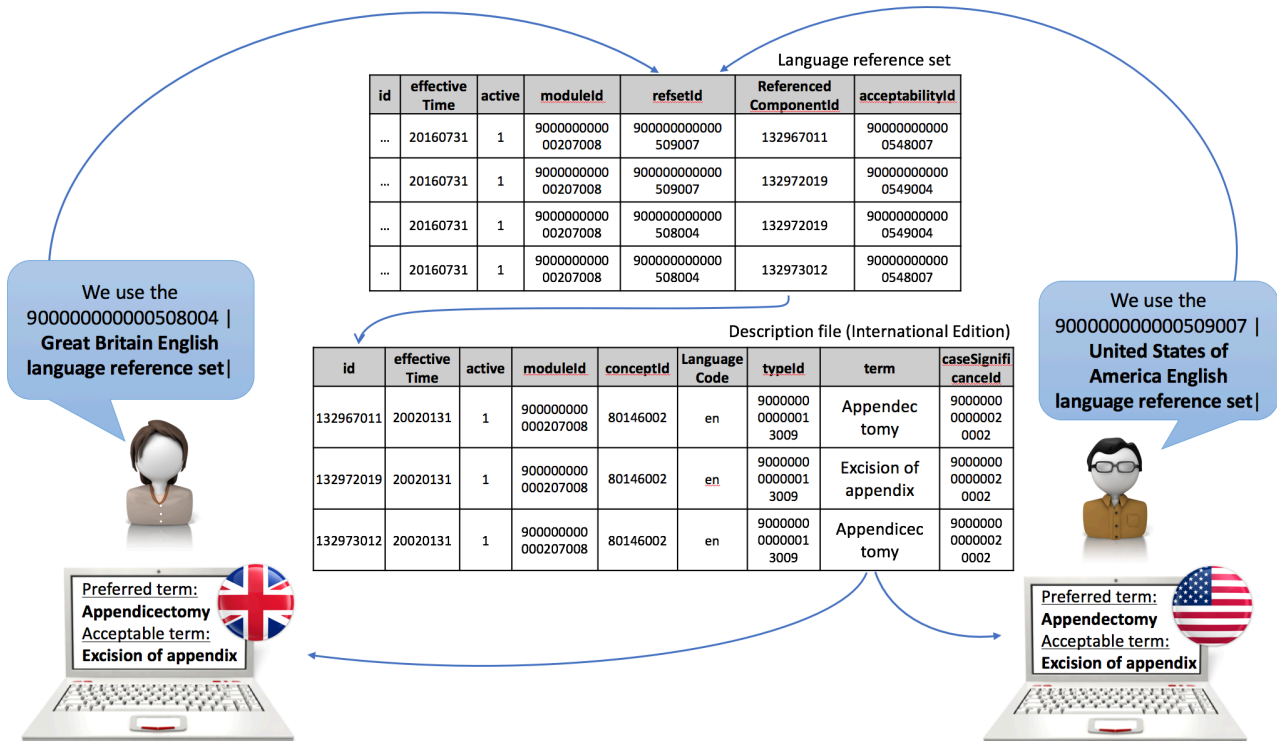
Figure 3.2.5.2-2: Use of language reference sets to specify preferred and acceptable descriptions for specific contexts.

Find more information here: 4.2.1 Language Reference Sets

Benefits of using SNOMED CT as Interface Terminology

The benefits of directly using the terms from SNOMED CT on the user interface are that

- no mapping is required from the clinical phrases to the clinical meanings. The design of SNOMED CT already include human-readable representations of concepts
- the clinical intent of a selection can be confirmed if required using other terms linked to the same concept (such as the 'Preferred Term')
- the equivalence of the terms to the clinical meaning is ensured by using quality authoring processes
- the higher quality of SNOMED CT coding can consequently lead to higher quality analytics results
- there are standard mechanisms provided by SNOMED CT for distinguishing acceptable and preferred terms in different clinical contexts
- where appropriate, the standardization of preferred terms can improve patient safety (in areas such as medication management)

Considerations

- This approach may require a transition of the user experience. However, it should be noted that new descriptions may be added to SNOMED CT to meet the expectations of the users.
- Subsets need to be created and maintained to support users in searching for and recording the appropriate SNOMED CT concepts.

Using a Separate Interface Terminology

An implementer who is motivated to introduce SNOMED CT records, but who is also keen to keep using an existing interface terminology, may choose to map between the interface terminology and SNOMED CT to enable that SNOMED CT is used for storage. Using this approach each item in the interface terminology is bound (or mapped) to an appropriate SNOMED CT concept. When the interface term is selected, the identifier of the bound SNOMED CT concept is stored in the record. It is important when an interface terminology is being used that the mapping to SNOMED CT is of sufficient quality (ideally equivalent) to support the use cases for which the data will be used. Using an interface terminology, for example, may be useful for structured data entry, where only part of the meaning is represented by the selected term, and the rest by the surrounding interface context. A simple map reference set can be used to represent the map between the interface terminology and SNOMED CT, in the case where there are a 1:1 map between each term in the interface terminology and SNOMED CT concepts.

3.2.6. Maintenance and Management

Managing the components used to represent the large number of clinical data entries in EHRs is an important part of the work related to maintaining the integrity and accessibility of health information. Like SNOMED CT components, reference sets are supported by a robust versioning mechanism that allows historically consistent views of SNOMED CT components and derivatives. This allows reference sets to be used to specify changes in use of concepts and descriptions in different parts of an EHR.

For more detailed use case examples, please refer to the following sections:

- 3.2.6.1. Constrain Value Sets
- 3.2.6.2. Managing Value Sets
- 3.2.6.3. Managing Component Inactivation

3.2.6.1. Constrain Value Sets

Most health records are designed and developed using one or more information models, which describe the information that is collected, stored, communicated and displayed. Some information models are designed for a specific proprietary system, while others are based on a common health information standard. Irrespective of the purpose, design and representation of the information models, the use of clinical terminology is an important part of making the models complete, meaningful and useful. Hence, a consistent approach to the interface between structural elements and terminological representations of information is required to support reliable interpretation of the meaning. Subsets of SNOMED CT components can function as value sets for any health-related information model to enable well-defined, unambiguous models of meaning.

As shown in the diagram below simple reference sets can be used to represent the subsets of SNOMED CT components to be populated as value sets within the relevant information models.
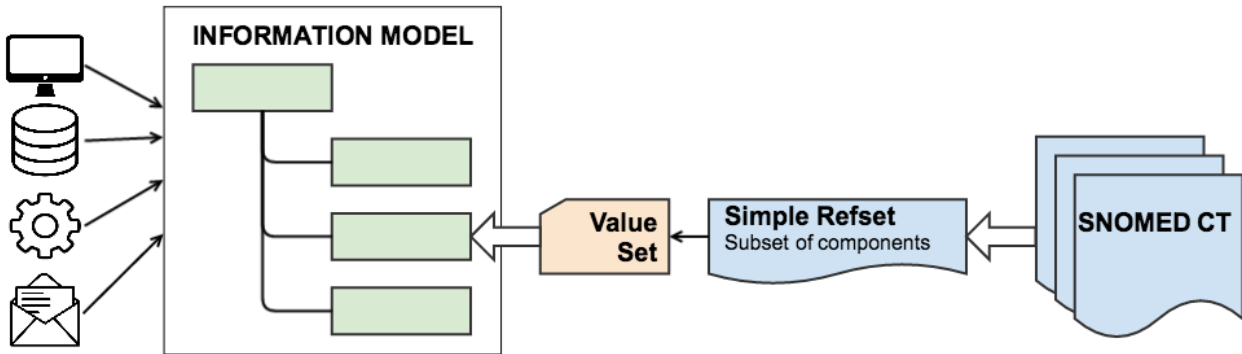
Figure 3.2.6.1-1: Relation between SNOMED CT reference sets, value sets and information models

3.2.6.2. Managing Value Sets

EHR systems will typically utilize a range of different value sets to be used in different places of the system. Representing these value sets using the same terminology will support the comparison of data captured in different contexts. Additionally, using SNOMED CT to represent items value sets instead of locally defined terms enables effective management and overview of information, and helps to mitigate challenges related to redundancy and ambiguity.

Simple reference sets can be used to represent extensionally defined subsets of SNOMED CT components, whereas the query specification reference set are useful for representing the intensional definition of SNOMED CT subsets. In the query specification reference set, the expression constraints can be used to represent the query used for defining the set. This means that the query specification reference set can be used to manage the intensional definition of SNOMED CT subsets that function as value sets, which is illustrated below.
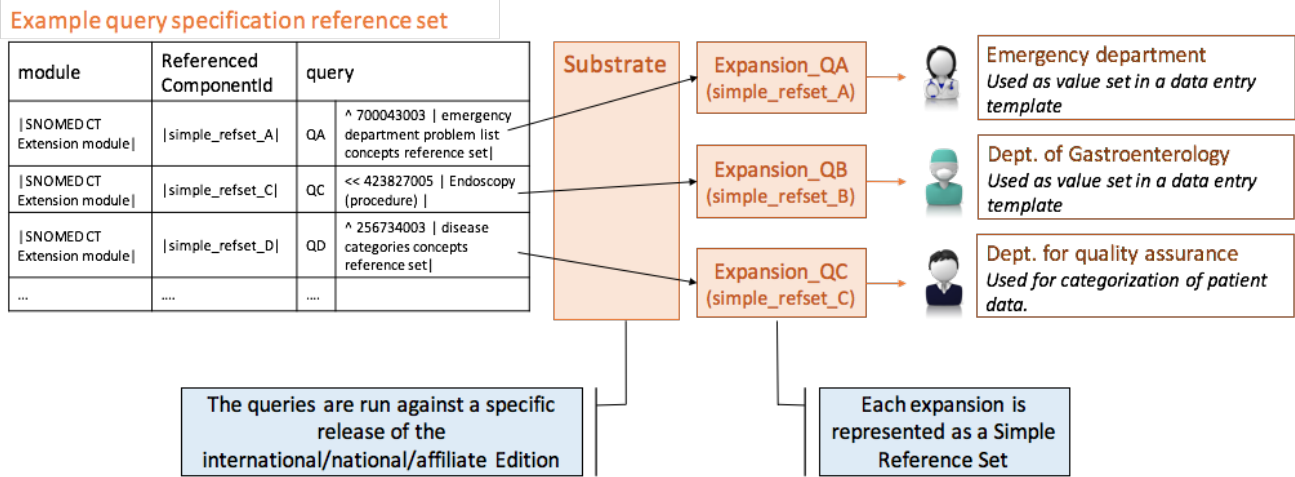
Figure 3.2.6.2-1: Query specification reference set used for generating value sets for different organisational units

3.2.6.3. Managing Component Inactivation

When a component that is a member of a reference set is inactivated, the person maintaining the reference set needs to decide whether a change is required. This depends on the intended use of the reference set being maintained. In the case of a reference set that is being used to constrain data entry, inactive concepts need to be removed from the reference set or replaced by an appropriate active concept. In other reference sets it may be permissible, or even required, to retain an inactive concept in a reference set (for example if a reference set is used in the criteria for a report which may be applied to historical data).

The following figure and subsequent description introduce the overall process for identifying inactive components and using released reference sets to determine reasons for inactivation and potential replacements.

| 1.
Identify components that have been inactivated since last release of SNOMED CT | 2.
Identify reasons for inactivation in the relevant \| Concept inactivation indicator attribute value reference set \| | 3.
Identify replacements using the appropriate \| Historical association reference set \| | 4.
Determine approach to manage the inactivation |
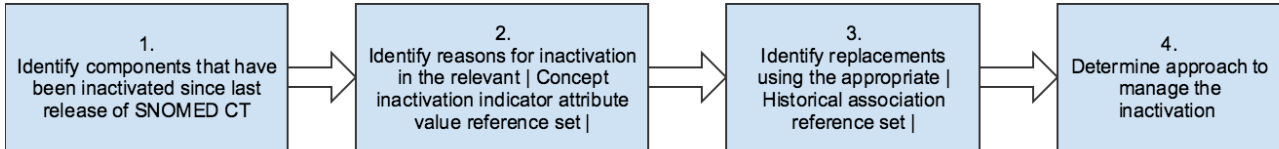
Figure 3.2.6.3-1: Process of determine reasons for inactivation and alternative replacements

1. One approach to identify the components that have been inactivated since the last release, is to compare snapshot of previous release with current delta release.
2. The reasons for inactivation can be looked up in the 900000000000480006 |Attribute value type reference set| for the particular component type, for example the 900000000000489007 |Concept inactivation indicator attribute value reference set| . The reason is represented by the value of the valueId attribute. For further information, see 3.2.6.3.1. Representing Reasons for Component Inactivation.
3. The possible replacements for the components that have been inactivated can be determined in the appropriate 900000000000 522004 |Historical association reference set| . For further information, see 3.2.6.3.2. Representing Historical Associations.
4. The preferred approach to manage inactivation of a component depends on the situation and the use of the reference set. Howe ver, a typical approach would be to update the reference set to apply the replacement concept instead of the inactivated concept. Please note, that some changes to the reference set may require additional updates to be performed to ensure correct use, see 6.6.2. Managing Reference Set Changes.

3.2.6.3.1. Representing Reasons for Component Inactivation

In the International Edition of SNOMED CT three reference sets of the type 900000000000480006 |Attribute value type reference set| are used to indicate the reason why components have been inactivated. These are:

- 900000000000489007 |Concept inactivation indicator attribute value reference set (foundation metadata concept)|
- 900000000000490003 |Description inactivation indicator attribute value reference set (foundation metadata concept)|
- 900000000000547002 |Relationship inactivation indicator attribute value reference set (foundation metadata concept)|

For example, if the inactivated component is marked as 900000000000484002 |Ambiguous| , this will typically mean that there will be more than one component replacing the inactivated component. The possible replacements will be available in the 900000000000523009 |POSSIBLY EQUIVALENT TO association reference set (foundation metadata concept)| .

Another example is concepts that are inactivated because they were 900000000000482003 |Duplicate| . This is used if two concepts represent the same meaning, because then one of those concepts must be inactivated. The equivalent concept will be represented in the 900000000000527005 |SAME AS association reference set (foundation metadata concept)| . Please refer to 3.2.6.3.2. Representing Historical Associations to see which reference sets relate to the various reasons for inactivation.

3.2.6.3.2. Representing Historical Associations

When a new version of SNOMED CT is released this may include changes to the content of the terminology. Components may have been added, inactivated or changed as described in 11.4 Versioning. As part of the terminology maintenance process, it may be appropriate to evaluate both the Representing Reasons for Component Inactivation and to determine appropriate alternatives.

The International Edition of SNOMED CT distributes a set of reference sets that record the reason that each inactive component was inactivated. These are referred to as "historical association" reference sets (see 4.2.3 Historical Association Reference Sets for more details). There is one historical association reference set for each type of historical association as shown in the table below.

Table 3.2.6.3.2-1: Association reference set types in the International Release of SNOMED CT

Association reference set	Descriptions
900000000000523009 |POSSIBLY EQUIVALENT TO association reference set|	Applies to a concept that is ambiguous. The targetComponent is an active concept that represents one of the possible meanings of the inactive concept . Multiple rows are used to refer to each of the possible meanings of the ambiguous concept.
900000000000524003 |MOVED TO association reference set|	Applies to a component that has been moved to (or are pending a move to) another namespace. The targetComponent identifies the target namespace (not the new component).
900000000000525002 |MOVED FROM association reference set|	Applies to a component that has been moved to this namespace from another namespace. The targetComponent identifies the original component Identifier in its previous namespace.
900000000000526001 |REPLACED BY association reference set|	Applies to an erroneous, obsolete and other inactive component for which there is a single active replacement. The targetComponent identifies the active component that replaces this component.
900000000000527005 |SAME AS association reference set|	Applies to a component that is a duplicate. The targetComponent identifies the active component that this component duplicates.
900000000000528000 |WAS A association reference set|	Links an inactive classification concept such as "not otherwise specified" or "otherwise specified" with the active concept that was formerly its most proximal supertype.
900000000000529008 |SIMILAR TO association reference set|	(not used currently)
900000000000530003 |ALTERNATIVE association reference set|	Links an inactive classification concept derived from ICD-9 Chapter XVI "Symptoms signs and ill-defined conditions" with the most similar active concept.
900000000000531004 |REFERS TO concept association reference set|	Applies to an inactive description which is inappropriate to the concept it is directly linked to but instead should refer to the concept referenced by the targetComponent.

The following table holds example entries for the \|Replaced by\| reference set. With this reference set is possible to automatically identify that the inactive concept 696005 \|Chronobiologic disorder\| should be replaced with the concept 387605007 \|Abnormal chronobiologic state\| . This type of reference set is particularly useful for ensuring consistent use of SNOMED CT over time. The reference set mechanism here provides an easy and standardized way of managing changes to coding or documentation practice over time.

Table 3.2.6.3.2-2: Sample content from 900000000000526001 | REPLACED BY association reference set |

referencedComponentId	referencedComponentId_term	targetComponentId	targetComponentId_term
100005	SNOMED RT Concept	138875005	SNOMED CT Concept
212002	Salmonella III arizonae 53:k:z	398450001	Salmonella IIIb 53:k:z
225005	Special care of patient with contagious disease	133895001	Care of patient with infectious disease
244003	Evans and Lloyd-Thomas syndrome	66659007	Normal variation in position
278009	Epidural injection of neurolytic substance, lumbar	17753007	Epidural injection of neurolytic solution, lumbar

558000	Other disorder of the neurohypophysis, NEC	72442006	Disorder of posterior pituitary
659001	Peptostreptococcus anaerobius	413524006	Anaerococcus tretradius
696005	Chronobiologic disorder	387605007	Abnormal chronobiologic state
700002	Salmonella III arizonae 50:z4,z23,z32:--	404619004	Salmonella IIIa 50:z4,z23,z32:-
822000	Salmonella arizonae 53:z4,z23:--	13998005	Salmonella IV 53:z4,z23:--

4. Reference Set Design

As the reference set mechanism is developed to support a range of different purposes, and therefore it has a high degree of flexibility and extensibility. As evidence of the flexibility in the design, some reference set types use a unique set of attributes in addition to the common reference set attributes. The additional attributes have been specified to meet the requirements for use of that particular type of reference set. The extendibility of the design allows reference sets to be customized to suite specific or local requirements. However, to support distribution, sharing and use of a reference set, it is important that the specific reference set design is consistently specified and represented. Thus, all attributes specific for a particular reference set type must be represented in a form which allows consumers of a reference set to validate and interpret the reference set.

The figure below provides an overview of some of the terms that are central for understanding the reference set design and how reference sets are specified and represented. Each of these terms will be elaborated in the following pages.
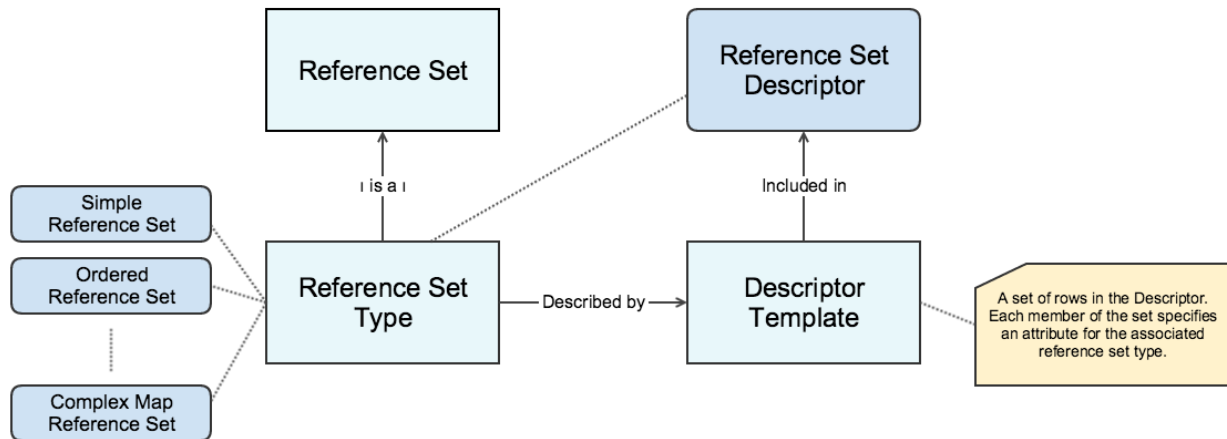
Figure 4-1: Relation between reference set types and descriptor templates

4.1. Reference Set Types and Descriptors

SNOMED International specifies a set of reference set types, which describes its own specific properties. This means that reference sets that are developed to conform to a specified pattern will have the same release file format as other reference sets of the same type.

The pattern of a specific reference set type is described by a Descriptor Template. This means that the Descriptor Template is represented by a set of members of the 900000000000456007 |Reference set descriptor reference set (foundation metadata concept)| . All Descriptor Templates present in the international Release of SNOMED CT can be found in the reference set Descriptor.

All reference sets that are released as part of the International Edition or from a National Release Center will have an associated Descriptor Template for the reference set. Where using a reference set for which a Descriptor Template has not been created, and additional information about the reference set is needed, the Descriptor Template of the closest ancestor of the concept describing the reference set that does have a Descriptor Template may be used. This means that a reference sets with no specified Descriptor Template inherits the Template from its supertype.

An organization that releases reference sets should only release them without Descriptor Templates if the reference set follows a predefined pattern or if it is sure that its consumers do not require the information held within the Descriptor Template. You should note that Descriptor Templates are optional for other organizations, besides from SNOMED International, that create reference sets that do not follow a predefined pattern. However, we strongly recommend to specify the reference set descriptor template in the reference set Descriptor, to support automatic processing, validation and sharing of the reference sets.The diagram below illustrates the different reference set types and highlight some of the specific reference sets that are included in the International Edition of SNOMED CT.
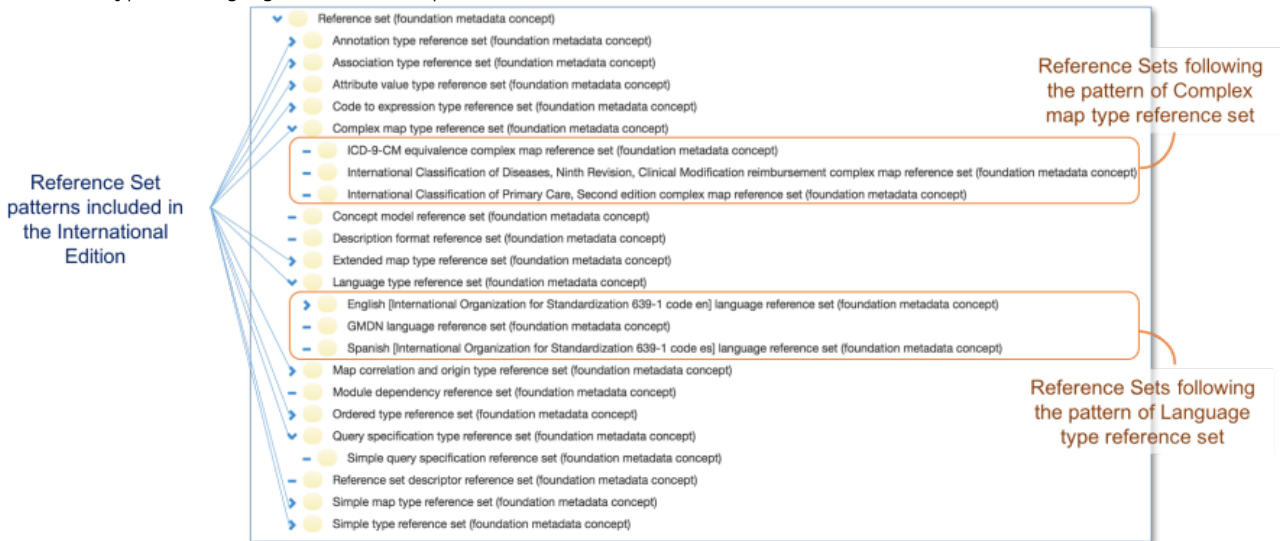
Figure 4.1-1: Reference set types and reference sets included in the International Edition of SNOMED CT

4.1.1. Reference Set Descriptor

The |reference set descriptor reference set| is a reference set used to specify the format of all reference sets included in a release. The data type and meaning of the referenced component and each additional field within each reference set is described by this reference set. More specifically, the reference set descriptor is used to define:

1. The order of appearance of additional attributes (other than those mandatory for all reference sets). The AttributeOrder attribute
2. The name and purpose of the additional attributes. The attributeDescription attribute
3. The data types for the additional attributes. The attributeType attribute

The table below shows an excerpt from the Reference Set Descriptor, to illustrate how the attributes of the predefined reference set types are specified consistently.

Table 4.1.1-1: Sample of the | reference set descriptor reference set (foundation metadata concept) | (human-readable view with some attributes omitted for brevity)

| ...	referencedComponentId	referencedComponentId_term	attributeDescription	attributeDescription_term	attributeType	attributeType_term
	446609009	Simple type reference set	449608002	Referenced component	900000000000461000	Concept type component
	447258008	Ordered type reference set	449608002	Referenced component	900000000000460000	Component type
	447258008	Ordered type reference set	447255006	Priority order reference set attribute	900000000000478000	Unsigned integer
	447258008	Ordered type reference set	447257003	"Linked to" reference set attribute	900000000000460000	Component type
	900000000000480006	Attribute value type	449608002	Referenced component	900000000000460000	Component type
	900000000000480006	Attribute value type	900000000000491004	Attribute value	900000000000461000	Concept type component
	900000000000496009	Simple map	900000000000500006	Map source concept	900000000000461000	Concept type component
	900000000000496009	Simple map	900000000000499002	Scheme value	900000000000465000	String
	900000000000506000	Language type	900000000000510002	Description in dialect	900000000000462000	Description type component
	900000000000506000	Language type	900000000000511003	Acceptability	900000000000461000	Concept type component
	900000000000512005	Query specification type reference set	900000000000514006	Generated reference set	900000000000461000	Concept type component
	900000000000512005	Query specification type reference set	900000000000515007	Query	900000000000465000	String
	900000000000516008	Annotation type	900000000000518009	Annotated component	900000000000461000	Concept type component
	900000000000516008	Annotation type	900000000000519001	Annotation	900000000000465000	String
	900000000000521006	Association type	900000000000532006	Association source component	900000000000460000	Component type
	900000000000521006	Association type	900000000000533001	Association target component	900000000000460000	Component type

Specializing a Descriptor Template

When a new reference set is created in an extension, it must (by default) conform to the reference set descriptors of its closest supertype (if one exists). Creation of a description template for a new reference set is optional in an extension, if the reference set has a supertype which itself has a descriptor template.

However, it is possible for a reference set to specialize the descriptor template of its supertype, by creating a copy and replacing the |Attribute description| or |Attribute type| values with a subtype.

For example, if a reference set has a descriptor, in which the |Attribute description| = |Referenced component| and the |Attribute type| = |Component type| , then the reference set's subtype may replace this descriptor with one in which the |Attribute description| = |Component type| and the |Attribute type| = |Concept type component| (since |Concept type component| is a subtype of |Component type|). In this way, the reference set descriptors may be specialized for use by the subtypes of the reference set .

4.1.2. General Functionality and Specific Use

All reference set patterns provide general functionality which is enabled by the attributes and data type constraints specified for that particular pattern.

For example, the general functionality of an association reference set is to represent a set of unordered associations of a particular type between SNOMED CT components. This general functionality may be sufficient to fulfill a range of different requirements. It may be used to associate inactive components with active concepts which can be used as

suitable replacements for the inactive concept. This use may be important for maintenance when a single reference set is used in a range of locations, and it is required to ensure consistent use of alternatives when content is inactivated. Another use of the same pattern may be to associate findings and procedures, which enables a simple form of conditional documentation support, for example, when a particular finding has been recorded, these are the procedures which may be appropriate.

Pattern	Simple type Reference Set	Ordered Reference Set	Association Reference Set	...
General functionality	A set of references to SNOMED CT components	A set of references to SNOMED CT components with a specified priority ordering	A set of unordered associations of a particular type between SNOMED CT components.	...
Specific uses	• Constrain search and data entry • Categorization of clinical data	• Order search results and data entry options • Specify alternative hierarchical views for navigation	• Associate inactive concepts with active concepts as potential replacements	...

Figure 4.1.2-1: Illustration of the general functionality and specific uses of selected types of reference sets

When deciding what reference set to develop it is therefore important to be aware of what requirements there are for use of that particular reference set, in order to decide on a pattern, which reflect the general functionality that meet a specific usage.

The name of a reference set pattern also reflects the general functionality of the pattern. reference sets that are developed following a specified pattern will be assigned a description including both a name describing that particular reference set and a term representing the pattern of the reference set. For further information, see 4.1.3. Naming Conventions for Reference Sets.

4.2. Common Reference Set Format

Reference sets are made available as release files that represent database tables. Each row of the table represents a member of the reference set [1] Each reference set member is represented using the set of attributes shown in the figure below. These common attributes include the four general identification and versioning attributes (shared with components) and two specific reference set attributes. This common set of attributes is sufficient to represent a versioned subset of SNOMED CT components.

As mentioned earlier, reference sets are used for purposes that go beyond subset representation. Different types of reference set are defined for each specific purpose. The definition of each reference set type specifies additional attributes and the way these are used. In these cases, each reference set member is represented by the combination of the common and type specific attributes.
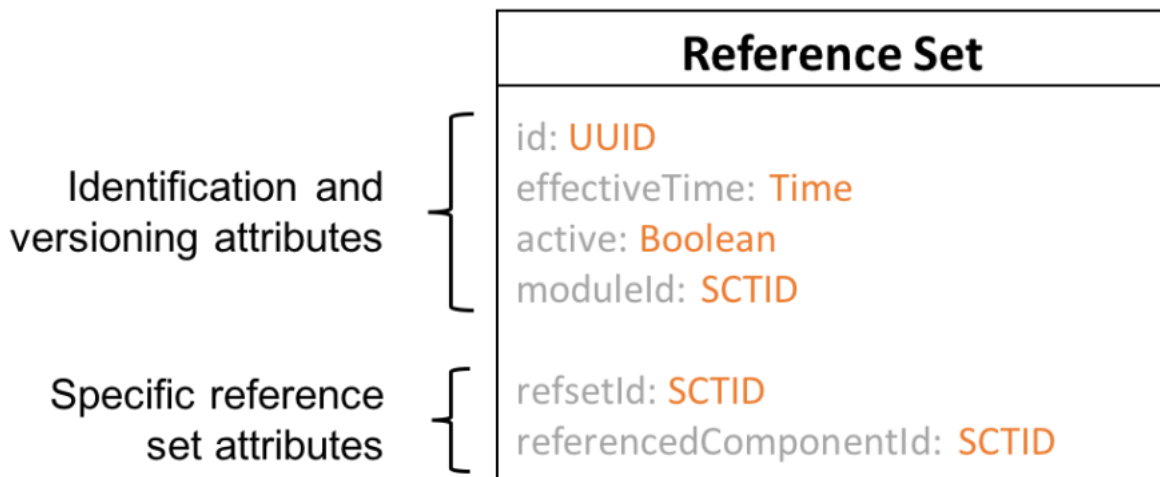
Figure 4.2-1: Attributes used in all types of SNOMED CT reference sets

Identification and Versioning Attributes

The reference sets have the same four initial attributes as the content components concepts, descriptions and relationships. These attributes are used to support identification, versioning and modularization.

Table 4.2-1: Overview and description of the general component attributes

Attribute	Description
id	The id attribute of the reference set uses the data type Universally Unique Identifier (UUID) to provide a globally unique identifier for each member of the reference set. This is different from the core Components of SNOMED CT, which use the data type SCTID. The UUIDs are 128-bit unsigned integers. Their unique values are generated by widely available algorithms and not part of the SCTID namespace. This avoids the need to track issuing of identifiers for thousands of reference set rows that are needed for some reference sets.
effectiveTime	The effectiveTime attribute uses the data type Time to specify the date on which the specific version of the component was released.
active	The active attribute uses the Boolean data type to specify whether or not the specific version of the component is active.
moduleId	The moduleId attribute identifies the module in which the component is currently being maintained.

The three attributes, id, effectiveTime and active are used together to version each component.

Specific Reference Set Attributes

There are also two attributes that are shared by all reference set types, but not with the content components (concepts, descriptions and relationships). These attributes are called the refsetId and the referencedComponentId.

Table 4.2-2: Overview and description of the Specific reference set attributes. These attributes are present for all types of reference sets, but some reference sets have more attributes

Attribute	Description
refsetId	The refsetId uses a SCTID to identify the reference set. The attribute refers to a concept and the concept's associated descriptions names the reference set. The concept is also a subtype to the concept that represents the type of reference set. One example of concept referenced by the refsetId is the concept 447566000 which description	Virtual medicinal product simple reference set	is the name of the reference set.
referencedComponentId	The referencedComponentId identifies a component referenced by the reference set. A reference set of type simple reference set represents a subset of SNOMED CT components. For this reference set type the referenced components are the subset members. One example of a component included in the 447 566000 |Virtual medicinal product simple reference set| is	Warfarin sodium 5mg tablet	

Reference Set Member and Referenced Component

The two terms "reference set member" and "referenced component" are sometimes used interchangeably, but when working with reference sets it is important to be aware of the distinct meaning of both of these terms, which should not be confused.
A reference set member is simply a single row in a specific reference set. Each member therefore includes the identifier of the member, the membership versioning information, the identifier of the component that is referenced by that row and all the other information that is recorded in a row of the reference set. In contrast, the 'referenced component' is the concept, description or relationship whose identifier appears in the referencedComponentId of the reference set.

Table 4.2-3: The referenced component is the value of the referencedComponentId attribute

Sample content from 447565001 |Virtual therapeutic moiety simple reference set| .	
refsetId	referencedComponentId (Referenced component)
447565001 |Virtual therapeutic moiety simple reference set|	211009 |Norethandrolone preparation|
447565001 |Virtual therapeutic moiety simple reference set|	302007 |Spiramycin|
447565001 |Virtual therapeutic moiety simple reference set|	449005 |Penicillin G procaine|
447565001 |Virtual therapeutic moiety simple reference set|	544002 |Melphalan|
447565001 |Virtual therapeutic moiety simple reference set|	669007 |Vaccinia virus vaccine|
447565001 |Virtual therapeutic moiety simple reference set|	796001 |Digoxin|
447565001 |Virtual therapeutic moiety simple reference set|	847003 |D-thyroxine preparation|
447565001 |Virtual therapeutic moiety simple reference set|	922004 |Pralidoxime|
447565001 |Virtual therapeutic moiety simple reference set|	1039008 |Mercaptopurine|
447565001 |Virtual therapeutic moiety simple reference set|	1148001 |Ticarcillin|

⚠ To be more precise, each row of a reference set table represents a version of a member of a reference set because, like SNOMED CT components, reference set members can be revised or inactivated by adding new versions to subsequent versions of a release file.

4.2.1. Reference Set Identification

For practical reasons, a reference set needs to be identified and named, so it can be referred to unambiguously. In a reference set, the reference set is identified by the refsetId attribute, which refers to a concept.
This concept is a subtype of the concept 900000000000455006 |Reference set (foundation metadata concept)| .
As for other all concepts in SNOMED CT, descriptions and relationships are are added to enable human-readable representation of the concept, and to place the concept in the SNOMED CT hierarchy. The descriptions provide the name of the reference set, and the relationship refers to the concept representing the reference set type. This is illustrated in the diagram below where the concept with the id 49999999102 has the associated description "Infectious disease simple reference set", and the relationship places the concept as a subtype of the concept 446609009 |Simple type reference set (foundation metadata concept)| .

id	effective Time	active	moduleId	refsetId	referenced ComponentId
0000d6608041-005b-5108-b608-07891b210365	20140731	1	19999999103	49999999102	82272006
00008774f374-5f54-5234-8af1-8aca6e73b5c2	20140731	1	19999999103	49999999102	6142004
00006ac9-2f97-5ed4-b37e-84cb158c388d	20140731	1	19999999103	49999999102	55604004

49999999102

Infectious disease simple reference set (EXAMPLE ONLY)

|is a|
subtype of

446609009
Simple type reference set
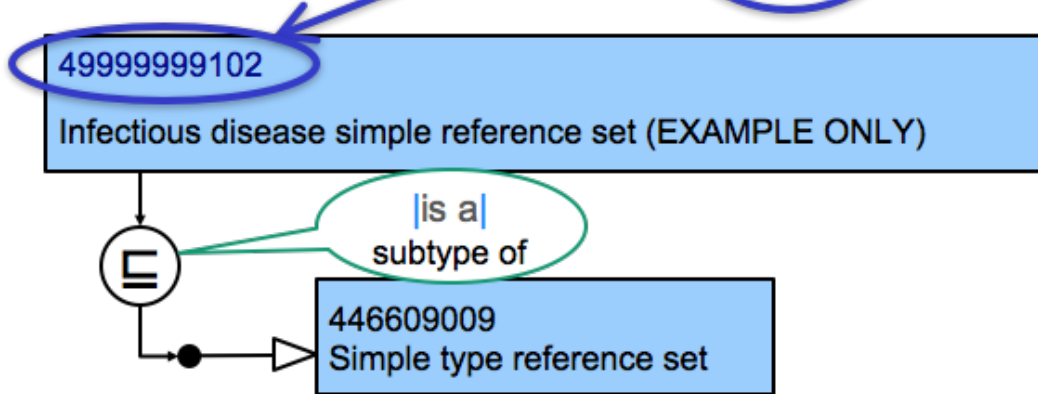
Figure 4.2.1-1: Reference sets are identified by SNOMED CT concepts

4.3. Pre-defined and Customized Reference Sets

SNOMED CT specifies a set of reference set types that follow specific patterns. However, different purposes and uses may impose different requirements. Therefore, flexibility is a core feature of the reference set mechanism. This flexibility is manifested by the fact that reference sets can be applied in ways that meet specific requirements; options include:

Option	Description
Create reference set based on existing pattern	Creating a simple reference set, an ordered reference set (or any other reference set) using the predefined pattern for this reference set type. Without modifying or constraining the structure of this pattern.
Create reference set based on existing pattern, however constraining the datatypes of specific attributes	Create a reference set using a predefined pattern, however specializing the descriptor template, by creating a copy and replacing the |Attribute description| or |Attribute type| values with a subtype.
Customize reference set by defining a new reference set pattern	Define a new reference set pattern based on new or existing reference set types, attributes, and attribute values. (Since you could use an existing type as a starting point.)

This means that if the pattern of existing reference set types is not sufficient for local requirements, then new reference set patterns can be specified and included in a national or local SNOMED CT Extension.

Customized Reference Sets

There might be situations where none of the defined reference set patterns are sufficient to fulfill the requirements for a specific use case for an implementing organization/project. In this case it is possible to develop customized reference sets, which include the exact attributes necessary to meet the requirements.

If an existing pattern almost meets the requirements a developing organization may want to simply add additional attributes to an existing reference set pattern. If none of the existing patterns can be adapted it is also possible to develop a new reference set pattern or a reference set that do not follow any specified pattern. The different approaches to customizing a reference set are illustrated in the figure below.
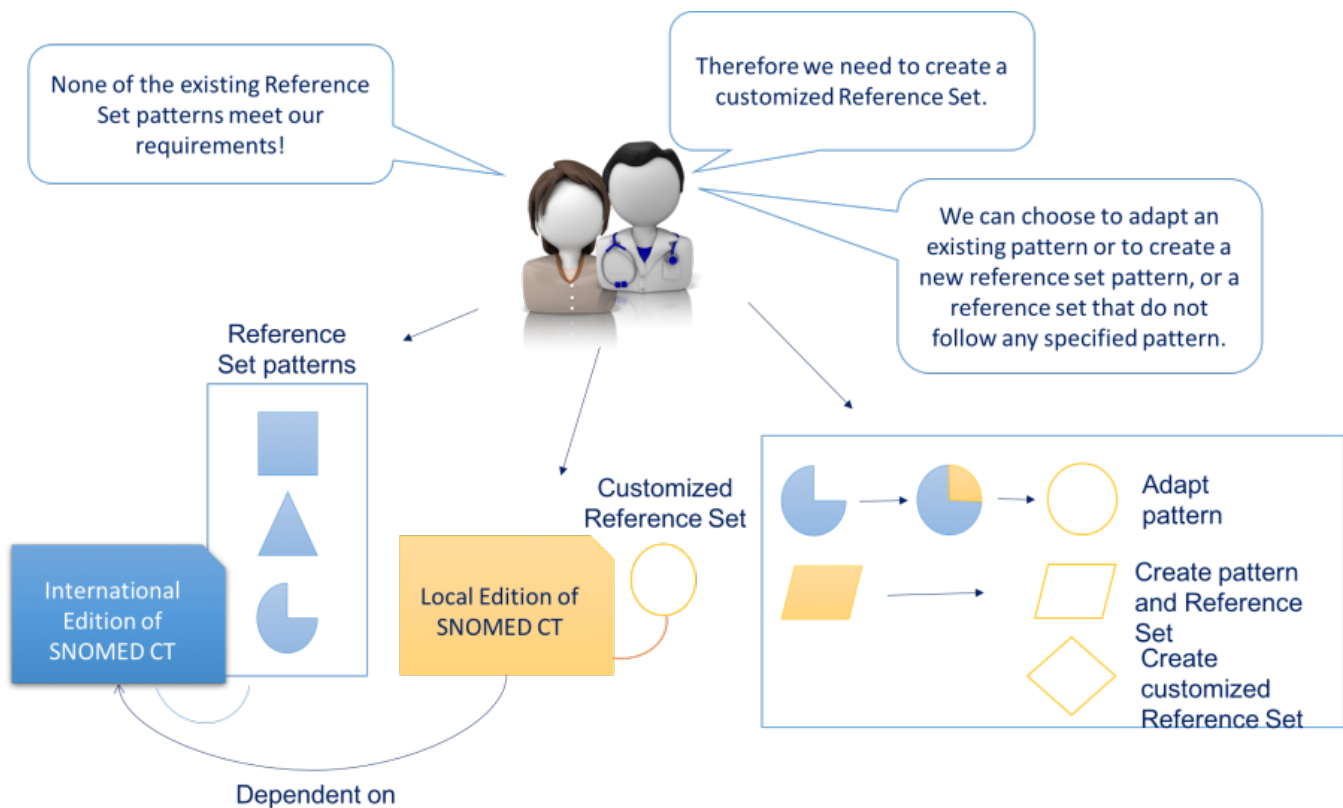
Figure 4.3-1: If no existing reference set pattern is sufficient to fulfill user requirements it is possible to create customized reference sets.

When creating customized reference sets, it is important to ensure that the structure and content of the reference set can be automatically processed and validated. Therefore, customized reference sets should be clearly specified, which can be done specifying the reference set pattern in the reference set descriptor. Specifying customized reference sets in the reference set descriptor enables users of the reference set to validate any reference set of this particular pattern against the specified definition.

5. Reference Set Types

SNOMED International specifies and distributes a range of different reference set types as part of the International release of SNOMED CT. These types of reference sets have been specified to provide a foundation for the more general requirements that may arise when implementing/enabling SNOMED CT in clinical information systems.

Each reference set type is described in a machine-readable form which follows a specific reference set pattern. An overview of the different types of reference sets is provided in Appendix 1: Overview of Reference Set Types and the technical specification for each of these types can be found in the Technical Implementation Guide, section 4. Reference Set Release Files Specification.
This guide will provide guidance to the usage of a selected set of reference set types. For details about the general functionality of each of the reference set types included in this guide, please use the links below :

- 5.1. Simple Reference Set
- 5.2. Query Specification Reference Set
- 5.3. Ordered Reference Set
- 5.4. Association Reference Set
- 5.5. Annotation Reference Set
- 5.6. Attribute Value Reference Set
- 5.7. Language Reference Set
- 5.8. Human-Readable Reference Set

These reference set types are included in this guide, as they relate to the practical use of SNOMED CT content in clinical information systems.
Other types of reference sets, such as the 4.2.8. Simple Map Reference Set and the 4.2.9. Complex and Extended Map Reference Sets set a re important for the co-existence of SNOMED CT and other code systems, but are addressed in other guides. Some reference sets are used to hold metadata, which is used for management or maintenance purposes, such as the historical association reference set and the 4.2.11. Module Dependency Reference Set. These reference set types will not be covered in this guide.

5.1. Simple Reference Set

The Simple reference set type represents an extensional definition of a subset of components. The components can be concepts, descriptions, relationships and reference sets. Therefore, the reference set contains a list of references to one or more components.

The components can be specified for inclusion or exclusion for a specified purpose. For example, the 450970008 |General Practice / Family Practice reference set| contains the concepts that are important for general practice and family practice medicine. In this section we will introduce the Simple type reference set format, the techniques for creation of simple reference sets and provide some examples of Simple reference sets and their usage.

Simple reference sets contain only the basic information needed to define a subset. As presented in the section about the reference set design, each member in a reference set has a referencedComponentId attribute, which is used to refer to the component that is a member of the reference set. Within the instances of this attribute, the individual references to components are stored. The diagram below illustrates an example of a simple reference set, and illustrates how the subset members are referenced in the referencedComponentId attribute.
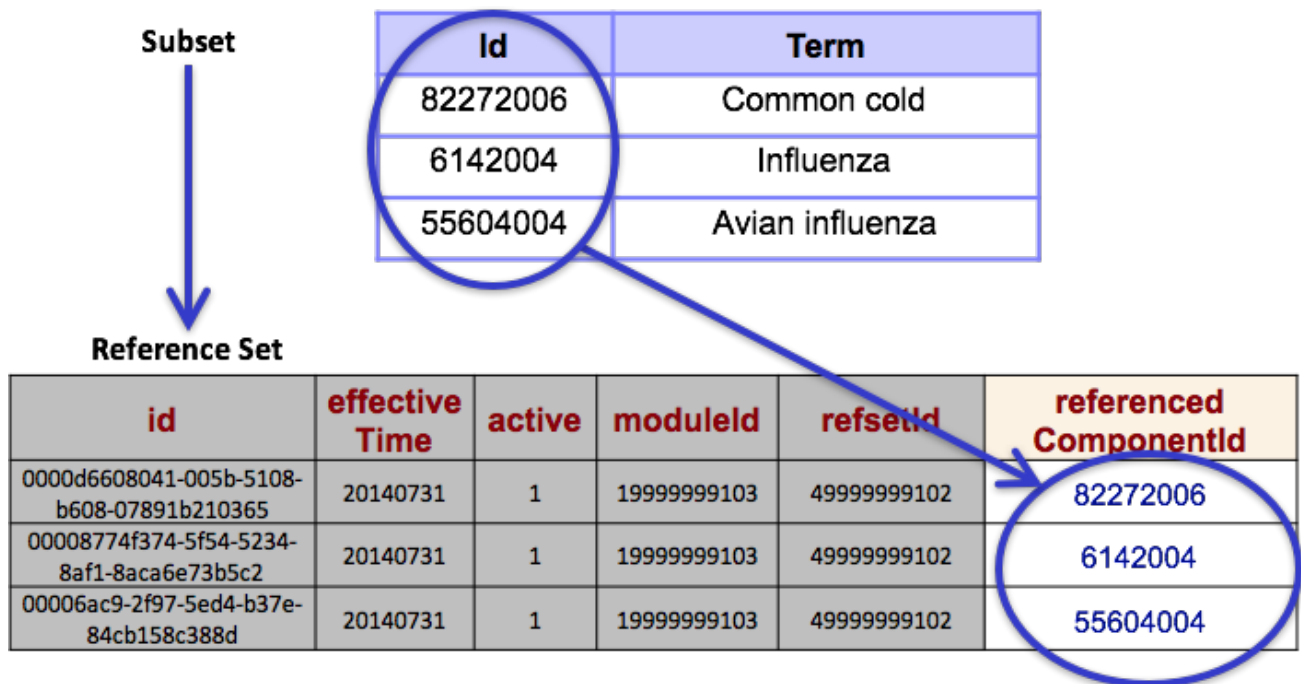

Figure 5.1-1: Simple reference set example

See specification: 4.2.1. Simple Reference Set

5.2. Query Specification Reference Set

The query specification type reference set is used to hold a series of queries used to represent the membership of a subset of SNOMED CT components.

A query contained in the reference set is run against the content of SNOMED CT (the substrate) to produce a subset of concepts, descriptions or relationships (the expansion). The query is referred to as an intensional definition of the subset. It can be run against new releases of SNOMED CT to generate an updated set of subset members. Additionally, the members of the expansion may also be represented in an enumerated form as a simple reference set. An enumerated representation of a subset is referred to as an extensional d efinition of a subset. The diagram below illustrates the relation between the query specification reference set and the simple reference set. The intensional definition is represented in the 'query' attribute, and the identifier for the expansion's simple reference set is represented by the concept referenced in the 'referencedComponentId' attribute.

The two tables below illustrate the relationship between the query specification type reference set and the simple reference set. The referencedComponentId in the query specification type reference set references a concept which represent the simple type reference set, which is used to hold the expansion of the intensional definition.

Table 5.2-1: Example Query Specification Type Reference Set

id	effectiveTime	active	moduleId	moduleId_term	refsetId	refsetId_term	referencedComponentId	referencedComponentId_te
<UUID>	20160131	1	19999999103	Example Extension Module	900000000000513000	Simple query specification reference set	739999999103	Route of administration simp reference set

Table 5.2-2: Sample from Resulting Simple Type Reference Set (only part of the expansion is shown here, as the concept| Route of administration value | has more subtypes than those shown here).

id	effectiveTime	active	moduleId	moduleId_term	refsetId	refsetId_term	referencedComponentId	referencedComponentId_term
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	420254004	Body cavity route (qualifier value)
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	419762003	Peritendinous route (qualifier value)
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	37161004	Rectal route (qualifier value)
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	419954003	Ileostomy route (qualifier value)
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	445754005	Intragingival route (qualifier value)
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	448077001	Intraepidermal route (qualifier value)
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	420163009	Esophagostomy route (qualifier value)
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	446442000	Transplacental route (qualifier value)
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	448491004	Intrajejunal route (qualifier value)
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	127490009	Gastrostomy route (qualifier value)
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	419243002	Transcervical route (qualifier value)

<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	...	...
<UUID>	20160131	1	19999999103	Example Extension Module	739999999103	Route of administration simple reference set	...	...

The Expression Constraint Language is the recommended language for specifying queries over SNOMED CT content. This language allows for a consistent and machine-readable representation of sets of clinical meanings. This means that a set of expression constraints can be used to specify the intensional definition of a range of subsets used in a particular context. The use cases for the subset generated from the queries within a query specification reference set are similar to the use cases for simple reference sets. However, the query specification reference set type can be used anywhere a set of queries needs to be managed.

Reference Set Specific Attributes

Field	Data type	Purpose
referencedComponentId	SCTID	The identifier (refsetId) of the reference set for which members are to be generated.
query	String	The serialised query that can be used to (re-)generate the reference set members. The SNOMED CT Query Language is a formal language for representing computable queries over SNOMED CT content.

See specification: 4.2.7. Query Specification Reference Set

5.3. Ordered Reference Set

The design of the ordered reference set supports three overall purposes:

1. To specify a sequential order of a subset of components
2. To specify prioritized groups within a subset of components
3. To define alternative hierarchies of components

Ordering

Ordered reference set can also be used to create a simple ordered list of components, i.e. a list that do not include any nesting, or groups. For ordered lists that do not require grouping or hierarchical arrangement the value of linkedToId should be the digit zero (0), as this attribute becomes irrelevant.

This type of ordered reference set can for example be used to prioritize the sort order of the descriptions with identical terms when they are displayed. It can also be used to specify the order of descriptions displayed in a simple pick list.
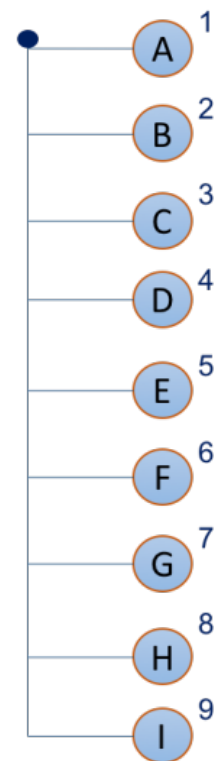
Figure 5.3-1: Ordered reference set with no groups

Prioritization

Prioritization is similar to order but multiple components may have the same rank. In this case the value of the order attribute specify a priority order for a group of components.

Ordered Reference Set → Groups

referencedComponentId	order	linkedToId
A	1	0
B	1	0
C	1	0
D	2	0
E	2	0
F	3	0
G	3	0
H	3	0
I	3	0
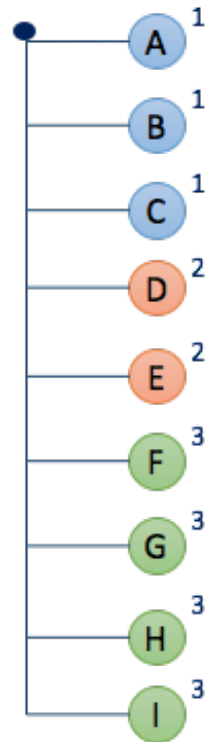
Figure 5.3-2: Ordered reference set with prioritized groups.

Alternative hierarchy

The diagram below Illustrates how the three attributes referencedComponentId, order and linkedToId are used to create an alternative hierarchical order of some of the concepts from the subtype hierarchy.
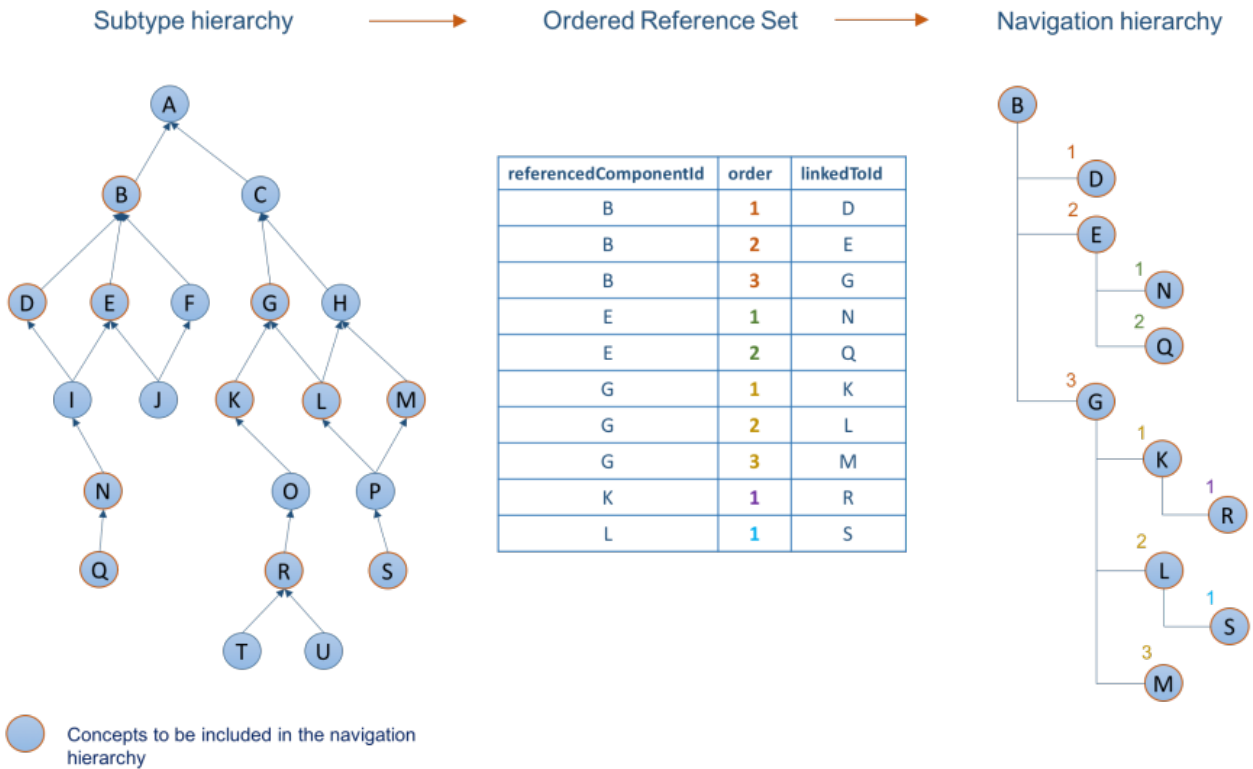
Figure 5.3-3: Ordered reference set example.

Reference Set Specific Attributes

Specific reference set attributes used to build an alternative hierarchical view of SNOMED CT

Attribute	Description
referencedComponentId	The identifier of a SNOMED CT component that is included in the ordered list of alternative hierarchy.
order	Specifies the sort order of the list. The list is ordered by applying an ascending sort of the order value. The value of order =1 represents the highest priority. A value of '0' is not allowed. Duplicate values are permitted and the sort order between two members with the same order value is not defined. If the linkedToId value is not 0, sorting occurs within subgroups that share the same linkedToId value.
linkedToId	The identifier of a SNOMED CT component that acts as a grouper or hierarchy node, collecting together a subgroup from within the list. This field either enables reference set member linked into a number of subgroups. These subgroups can be nested allowing representation of alternative hierarchies. To link members into a subgroup, all components in the same subgroup should reference the same component. This can either be a component that represents the name of that subgroup or the first member of the subgroup. In the latter case, the first row of each subgroup will contain the same identifier in referencedComponentId and linkedToId and with order =1. To link a number of children concepts to a single parent concept, one member record should exist per child, with the referencedComponentId field referencing the parent and this field referencing the child concept. The order field is then used to order the children concepts under the parent concept.

See specification: 4.2.2. Ordered Reference Set

5.4. Association Reference Set

The Association type reference set represents a set of unordered associations of a particular type between components, i.e. concepts, descriptions and relationships. The referencedComponentId references the source component of the association, whereas the reference set specific attribute, the targetComponentId, references the target component of the association.
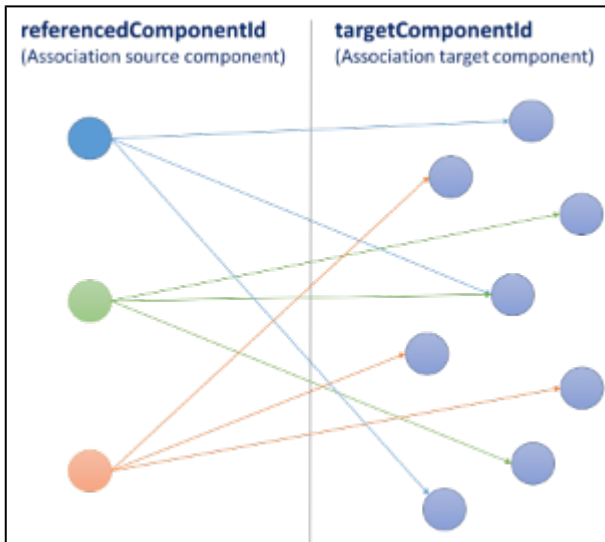
Figure 5.4-1: Illustration of the overall functionality of Association reference sets

Associations between components can be used to form groups of components, as illustrated in the diagram below. When using an association reference set to specify group, the sort order of the group members is not specified, as when using the Ordered type reference set.
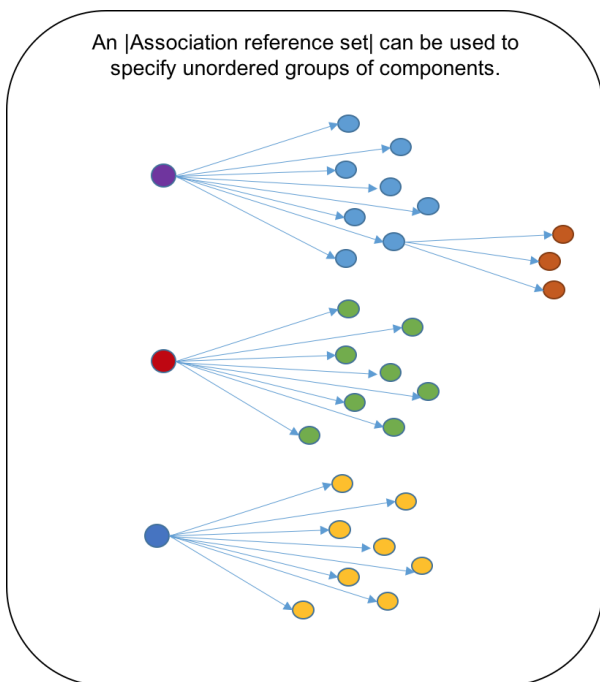
Figure 5.4-2: Using an |Association reference set|to specify groups of components

Reference Set Specific Attributes

Field	Data type	Purpose
referencedComponentId	SCTID	The identifier of the source component of the association.

targetComponentId	SCTID	The identifier of the target component of the association.

See specification: 4.2.5. Association Reference Set

5.5. Annotation Reference Set

|Annotation type reference set| allows text strings to be associated with components for any specified purpose. So, where the |Association type reference set| linked a SNOMED CT component to another SNOMED CT components, the |Annotation type reference set| allow a SNOMED CT component to be linked to a non-standardized string annotation.
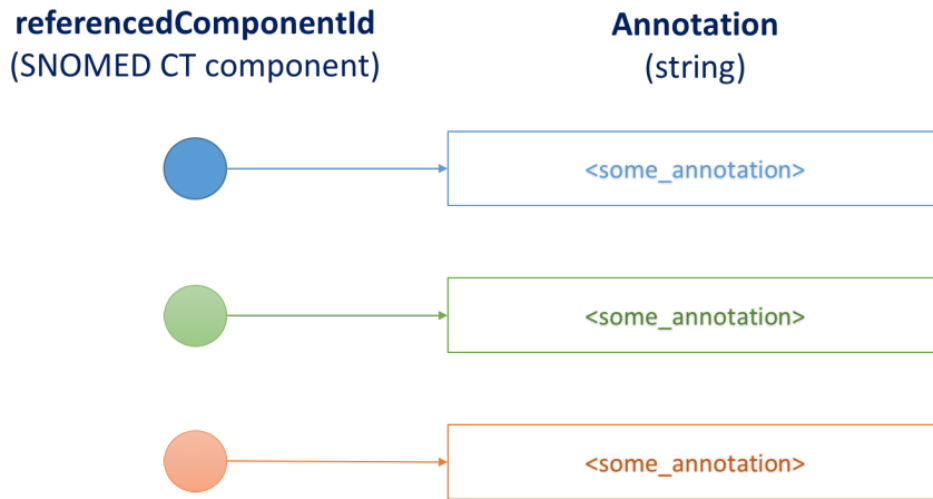
Figure 5.5-1: Illustration of the general functionality of Annotation reference sets of the general functionality of Annotation reference sets

Reference Specific Attributes

Besides from the 4 identification and versioning attributes, the annotation reference set type has following attributes.

Field	Data type	Purpose
referencedComponentId	SCTID	The identifier of the component to be annotated.
annotation	String	The text annotation to attach to the component identified by referencedComponentId.

See specification: 4.2.6. Annotation Reference Set

5.6. Attribute Value Reference Set

An attribute value reference set is a component reference set used to apply a tagged value to a SNOMED CT component. The pattern of the attribute value reference set is similar to that of the association reference set, despite the fact that the components referenced in the valueId attribute must be a subtype of the concept 900000000000491004 |Attribute value| , which include concepts such as:

An 900000000000480006 |Attribute value type reference set| allows a value from a specified range to be associated with a component. This type of reference set can be use for a range of purposes where there is a requirement to provide additional information about particular concepts, descriptions or relationships. In the International Edition of SNOMED CT an 900000000000480006 |Attribute value type reference set| is for example used to indicate the reason why components have been inactivated.

- 900000000000489007 |Concept inactivation indicator attribute value reference set (foundation metadata concept)|
- 900000000000490003 |Description inactivation indicator attribute value reference set (foundation metadata concept)|
- 900000000000547002 |Relationship inactivation indicator attribute value reference set (foundation metadata concept)|

Refset Specific Attributes

Field	Data type	Purpose
referencedComponentId	SCTID	A reference to the SNOMED CT component being tagged with a value.
valueId	SCTID	The tagged value applied to the referencedComponentId. A subtype of 900000000000491 004 |Attribute value| .

See specification: 4.2.3. Attribute Value Reference Set

5.7. Language Reference Set

Language reference sets are used to specify the acceptability and preference for using a particular description in a specific language, dialect or clinical context. To understand the importance of Language reference sets it is necessary to be aware of the characteristics of SNOMED CT Descriptions and their representation.

Descriptions and Language Reference Set

SNOMED CT contains descriptions and each description contains both a human-readable term and some information about the term. A description is used to give meaning to a concept and provide well-understood and standard ways of referring to a concept. So, each description is associated with a specific concept, but each concept is associated to several descriptions. Additionally, the descriptions are specific to a language or dialect.

All Descriptions are of a specific type and the most common description types are the Fully Specified Name and Synonym. The Description file holds descriptions that describe SNOMED CT concepts. The description file is released with the International Edition of SNOMED CT. National or Affiliate Editions may develop their own description file, e.g. to represent description in their own language and/or dialects.

The Language reference set is essential to enable the preferred terms to be identified for each concept. Language reference sets refers to Descriptions that is used in a particular language or dialect. For each Description referenced in the language reference set a value for the acceptability of the term associated with that Description is assigned. The values for the Acceptability attribute is represented by descendants of the concept 900000000000511003 |Acceptability| , which is placed in the foundation metadata subhierarchy of SNOMED CT.

Table 5.7-1: Available values for acceptability

Value	Description
900000000000548007	Preferred (foundation metadata concept)		The term associated with this description is the preferred description, of the specified Description.typeId, for the associated concept, in the language or dialect represented by this reference set. If the Description.typeId is synonym, this description is the preferred term. If the Description.typeId is fully specified name this description is the preferred fully specified name. For each concept there should be exactly one preferred description of each Description.typeId in each language reference set.
900000000000549004	Acceptable (foundation metadata concept)		The term associated with this description is acceptable for use in language or dialect represented by this reference set.

The following diagram shows an example of the description file included in the International Release and the US Language reference set, which is also distributed with the International Release of SNOMED CT. The language reference set states the acceptability of the descriptions, i.e. whether they are preferred or synonym. For each concept there may be any number of acceptable descriptions of each description type in each language reference set. The diagram below shows how the description file holds information about the description type, and the language reference set specify the acceptability of the descriptions.

Figure 5.7-1: Example of the description file included in the International Release and the US Language reference set

Even though a concept has several descriptions of the type 900000000000013009 |Synonym (core metadata concept)| related to it, the language reference set allows automatic identification of the preferred term . The language reference set does not contain any terms, as these are represented in the description file. Hence, it refers to descriptions by referencing the descriptionId, and the description release file is therefore a prerequisite for retrieving the actual term when applying the language reference set .

Table 5.7-2: Reference Set Specific Attributes

Field	Data type	Purpose
referencedComponentId	SCTID	The identifier of a description included in the language reference set.
acceptabilityId	SCTID	A subtype of 900000000000511003 |Acceptability| indicating whether the description is acceptable or preferred for use in the specified language or dialect .

See specification: 4.2.4. Language Reference Set

5.7.1. Language Reference Sets in National Editions

When SNOMED CT is translated into other languages, it requires the Extension to include a description release file containing the new descriptions with the translated and approved terms. It also requires a language reference set, which specifies the acceptability of these new descriptions. The language reference set necessary to distinguish the preferred synonym (preferred term) from the acceptable synonyms, and it is required for specifying the preferred Fully Specified Name (FSN) within that language. A concept may have more than one FSN, but only one of these may be marked as 'preferred' in a given language. A Language Reference Set is therefore used to specify which FSN description is preferred in each language or dialect.

Even when SNOMED CT is not translated into other languages, a Language Reference Set can be used within an Extension to specify which of the existing descriptions from the International Edition are preferred and accepted within the particular context where the Extension is applied. The table below illustrates that within a single SNOMED CT Edition, multiple language reference sets can be created.

Table 5.7.1-1: Language reference sets supported in the International Release and in the Canadian, Danish and Swedish Editions

Edition	Available Language reference sets
International Edition	▲ Reference set (foundation metadata concept) ▼ Language type reference set (foundation metadata concept) ▼ English [International Organization for Standardization 639-1 code en] language reference set (foundation metadata concept) – Great Britain English language reference set (foundation metadata concept) – United States of America English language reference set (foundation metadata concept) – GMDN language reference set (foundation metadata concept) – Spanish [International Organization for Standardization 639-1 code es] language reference set (foundation metadata concept)
Canadian EN Edition	▲ Reference set (foundation metadata concept) ▼ Language type reference set (foundation metadata concept) ▼ English [International Organization for Standardization 639-1 code en] language reference set (foundation metadata concept) – 🇨🇦 Canada English language reference set (foundation metadata concept) – Great Britain English language reference set (foundation metadata concept) – United States of America English language reference set (foundation metadata concept) ▼ 🇨🇦 French [International Organization for Standardization 639-1 code fr] language reference set (foundation metadata concept) – 🇨🇦 Canada French language reference set (foundation metadata concept) – GMDN language reference set (foundation metadata concept) – Spanish [International Organization for Standardization 639-1 code es] language reference set (foundation metadata concept)
Danish Edition	▲ Reference set (foundation metadata concept) ▼ Language type reference set (foundation metadata concept) – 🇩🇰 Danish language reference set (foundation metadata concept) ▼ English [International Organization for Standardization 639-1 code en] language reference set (foundation metadata concept) – Great Britain English language reference set (foundation metadata concept) – United States of America English language reference set (foundation metadata concept) – GMDN language reference set (foundation metadata concept) – Spanish [International Organization for Standardization 639-1 code es] language reference set (foundation metadata concept)
Swedish Edition	▼ Language type reference set (foundation metadata concept) ▼ English [International Organization for Standardization 639-1 code en] language reference set (foundation metadata concept) – Great Britain English language reference set (foundation metadata concept) – United States of America English language reference set (foundation metadata concept) – GMDN language reference set (foundation metadata concept) – Spanish [International Organization for Standardization 639-1 code es] language reference set (foundation metadata concept) – 🇸🇪 Swedish [International Organization for Standardization 639-1 code sv] language reference set (foundation metadata concept)

5.8. Human-Readable Reference Set

It is acknowledged that members and affiliates developing reference sets have the need to create human-readable versions of the reference sets. Human readable reference sets are used to allow reference sets to be inspected without requiring a software-service to resolve the id's referenced in the reference set.

Supporting the readability of reference sets is particularly important for:

- Review
  - I.e. Subject matter experts who are involved with reviewing the content of reference sets
- Educational and dissemination
  - I.e. when explaining the structure and content of reference sets

Purpose of Human Readable Reference Sets

SNOMED International specify a recommended format for human-readable reference sets to

- Avoid people reinventing and coming out with different solutions
- Allow a standard representation of reference sets shared as part of agreements with other standards bodies
- Allow tooling support (whether internally or externally developed)

Format

A column is added for the human-readable version of an attribute, with the addition of '_term' to the attribute name.

Examples of this include

- The corresponding human-readable column for the attribute refsetId is refsetId_term
- The corresponding human-readable column for the attribute referencedComponentId id referencedComponentId_term.
- The corresponding human-readable column for the attribute acceptabilityId is acceptabilityId_term

Table 5.8-1: An example of a standard format Language reference set

...	refsetId	referencedComponentId	acceptabilityId
	900000000000508004	42969009	900000000000548007
	900000000000508004	42969009	900000000000549004
	900000000000508004	80146002	900000000000549004
	900000000000508004	80146002	900000000000548007
	900000000000508004	271737000	900000000000548007
	900000000000508004	271737000	900000000000549004

Table 5.8-2: An example of a human-readable version of a Language reference set

...	refsetId	refsetId_term	referencedComponentId	referencedComponentId_term	acceptabilityId	acceptabilityId_term
	900000000000508004	GB English	42969009	Cauterisation of skin	900000000000548007	Preferred
	900000000000508004	GB English	42969009	Fulguration of subcutaneous tissue	900000000000549004	Acceptable
	900000000000508004	GB English	80146002	Excision of appendix	900000000000549004	Acceptable
	900000000000508004	GB English	80146002	Appendicectomy	900000000000548007	Preferred
	900000000000508004	GB English	271737000	Anaemia	900000000000548007	Preferred
	900000000000508004	GB English	271737000	Absolute anaemia	900000000000549004	Acceptable

Characteristics of the Term

When creating a human-readable refset it should be considered which term to include for each of the human-readable attributes. It is often preferable to include the preferred term from an appropriate dialect within the expressions to improve the human-readability of the attribute value. However, if it is required to disambiguate which hierarchy the concepts come from then the FSN can be used.

Human-readable Annotations to Reference Set Attributes and/or Members

Sometimes it may be relevant to add one or more columns to function as a placeholder for comments related to either specific attributes, or for the reference set member (row) as a whole. To enable that these columns can be automatically identified as non-standard specific, we recommend to use following naming convention for these columns.

The '_'-symbol is used to represent that this column represent comments, or information, which shouldn't be included when the reference set is used as part of a SNOMED CT implementation.

It is therefore important that fields holding comments are not used for purposes important for use of the reference set. If the column is central for the machine-processable version of the reference set, then the column should be specified as a reference-specific attribute in a customized reference set (4.3. Pre-defined and Customized Reference Sets) and identified in the reference set descriptor (4.1. Reference Set Types and Descriptors).

The '_'-symbol is used as a prefix in front of a self-determined string. I.e the string can be any term which make sense for the actual situation, e.g. '_comment' can be used to allow reviewers to provide a comment to the specific attribute, or '_alternatives' may be used to suggest alternative members or member values, which can be used for proposals as part of a review process while the refset is under development.

SNOMED International recommends two uses of the '_'-symbol:

- <attributename>_<string> is used to refer to a specific attribute value
- _<string> is used to refer to a whole reference set row, i.e a specific version of a reference set member.

Table 5.8-3: Example of comments related to specific attribute values

...	refsetId	refsetTerm	referencedComponentId	referencedComponentTerm	referencedComponentId_comment	referencedComponentId_alternatives	acceptabilityId	a
	900000000000508004	GB English	42969009	Cauterisation of skin	comment related to the concept | Cauterisation of skin |	suggested alternative(-s)	900000000000548007	Pre
	900000000000508004	GB English	42969009	Fulguration of subcutaneous tissue			900000000000549004	Ac
	900000000000508004	GB English	80146002	Excision of appendix	comment related to the concept | Excision of appendix |		900000000000549004	Ac
	900000000000508004	GB English	80146002	Appendicectomy		suggested alternative(-s)	900000000000548007	Pre
	900000000000508004	GB English	271737000	Anaemia	comment related to the concept | Anaemia |	suggested alternative(-s)	900000000000548007	Pre
	900000000000508004	GB English	271737000	Absolute anaemia	comment related to the concept | Absolute anaemia |		900000000000549004	Ac

Table 5.8-4: Example of comments related to a whole reference set member (a version of a refset member)

...	refsetId	refsetTerm	referencedComponentId	referencedComponentTerm	acceptabilityId	acceptabilityTerm	_comment
	900000000000508004	GB English	42969009	Cauterisation of skin	900000000000548007	Preferred	
	900000000000508004	GB English	42969009	Fulguration of subcutaneous tissue	900000000000549004	Acceptable	comment related to the specific row
	900000000000508004	GB English	80146002	Excision of appendix	900000000000549004	Acceptable	
	900000000000508004	GB English	80146002	Appendicectomy	900000000000548007	Preferred	
	900000000000508004	GB English	271737000	Anaemia	900000000000548007	Preferred	comment related to the specific row
	900000000000508004	GB English	271737000	Absolute anaemia	900000000000549004	Acceptable	

Considerations for Use of the Human-Readable Reference Set

For all uses of human-readable reference sets and use of additional annotations it is very important that users do not include information which affects the use of the reference set within the human-readable columns (columns with the '_' prefix). All information relevant for implementation, processing and processing of the reference set should be represented according to the standard, machine-processable format. Consequently, all columns with the '_'-prefix should be able to be ignored when the reference set is implemented.

6. Reference Set Development

General Development Process

The lifecycle of reference set work spans from the point where there is an initial idea related to a specific clinical information requirement based on SNOMED CT through reference set development to long term maintenance. At each stage it is essential that a defined process is followed which ensures the quality and validity of the product is maintained. The following figure provides an overview of the steps that are relevant to any project that either builds its own reference set, or adapts reference sets authored from elsewhere. The general process is similar to the process used when developing other information artefacts, and as for many development processes it is essential to be clear and specific about the requirements to be met, and to ensure that these requirements drive the design and implementation of the artefact. Related to reference set development this means that the process is not just about creating a reference set, but more about how to address the requirements, and this will often mean a combination of development of different reference sets.
The overview provided in following diagram is not intended to imply that a waterfall methodology should be used to develop a reference set. It may be delivered using agile or iterative project methodologies, with the reference set and supporting project documentation evolving as the project develops.
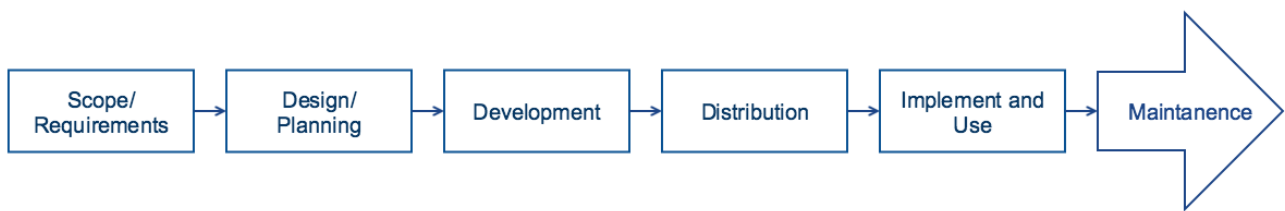

Figure 6-1: Reference set development process

6.1. Defining Requirements and Scope

The initial focus of the reference set development work is to establish a clear set of initial requirements, which may require further elaboration at a later date. The requirements should include an assessment of the release requirements for the reference set from a user perspective.

Defining the scope of the reference set involves answering a range of different questions related to each step of the development process. Following questions are relevant to consider when defining the scope of a reference set. Additional questions will probably be relevant for a specific development process.

- What is the main purpose of the reference set?
- Who are the users of the reference set?
- What is the scope of content of the reference set?
- Who is responsible for developing and maintaining the reference set?
- Should we develop a new reference set or does any existing reference set meet our requirements?

Once the scope of the reference set and the requirements has been identified, it is essential that an owner of the reference set is clearly identified (this may be an individual or a specific group). Furthermore, the requirements must be clearly documented, particularly the scope, purpose and use case(-s).

The owners of the reference set must then assess what content to be included in the reference set. This process may involve evaluating existing reference sets, or assessing what parts within the International Release of SNOMED CT (and potentially National and Affiliate Editions) will meet the stated requirements. At this stage the requirements for ensuring the content of the subset is quality assured should also be documented.

6.2. Design and Planning

Careful planning in the initial stage is crucial for successful completion of the reference set development and onwards use of the reference set. Following considerations should be included in any reference set development process.

Table 6.2-1: Considerations important when designing the reference set and planning the development process

Topic	Description
Resources	Determine the resources available for developing and maintenance of the reference set.
Dependencies	Dependencies of the reference set • What module and release of SNOMED CT is the reference set dependent on? • What other terminology and/or software artefact should the reference set function with? (see section on 6.2.2. Reference Sets and Information Models)
Reference set type	What type of reference set is required? • Is one of the existing reference set patterns sufficient? • Should a customized reference set be specified, and if so what additional attributes should be included?

People	Developing a new reference set requires a broad skillset, so it should be determined who will develop the reference set. Each of the steps related to developing reference sets can involve a range of people in different roles, a number of interdependent activities and a suite of related documents and artifacts. It is important to understand the characteristics of the people involved and their roles across the stages in the reference set lifecycle.
	
	• Some of the important skills required when developing reference sets include:
	  • clinical insight. To know what content is relevant for the specific purpose.
	  • sufficient knowledge on SNOMED CT. To support correct use of SNOMED CT, for example selection of appropriate concepts or defining the correct query.
	  • knowledge about the technical environment where the subset is going to function. In order to ensure an appropriate representation of the subset to be integrated in an IT-system.

Development approach and methods	What approach should be taken to develop the reference set? • Does any existing reference set meet the requirements for this new reference set or is there a need to build a new reference set from scratch? (see section on 6.2.1. Exploring and Evaluating Reference Sets)
Tools	What tools should be used to support reference set development, distribution and maintenance process? There will be a suite of tools provided which can support the creation and management of reference sets, including tools for: • Selection of reference set members, or definition of the query needed to retrieve reference members. Such tool will typically allow collaborative approaches to be pursued. • Support review by subject matter experts and others involved with creation and quality assurance • Managing requests for change through a reporting mechanism that will allow potential changes to the reference set to be identified. • Updating the reference set, e.g. when changes has been requested, or when new releases of the Editions on which the reference set is dependent upon is freed. • SNOMED International provides a range of services and tools to support working with SNOMED CT, and the suite of services continue to grow. Specifically, the SNOMED International reference set management tool, which is useful the creation and management of both intensionally and extensionally defined subsets.
Timeline	Determine when the reference set should be ready for routine use, i.e. schedule deadlines for the difference phases of the development process.
Quality assurance	What level of quality is required for the reference set? E.g. is the reference set used to represent safety critical information for patients, or is the reference set used to form broader categories of patient groups. It should be determined how the required level of quality is obtained and ensured throughout the development process.
Distribution	How will the reference set be distributed? Answering this question will depend very much on who the authors of the reference set is and where the reference set going to be used. If the reference set is used as a national reference set, i.e. a reference set which is used across a range of hospitals or organizations, distribution is likely to be done via a central service. Locally defined reference sets, e.g. used as part of configuring an EHR is likely to be managed and distributed from a local terminology storage environment and integrated directly to the system.
Integration	How, who and when to integrate the reference set into the environment where it is going to function? Decide on an approach to integration and determine what bindings and/or integration tasks to be performed.
Maintenance	Will new versions of the reference set be integrated regularly or as in integrated part of a greater maintenance process together with other software/terminology artefacts?

6.2.1. Exploring and Evaluating Reference Sets

As part of the design and planning of a new reference set may it may be enumerative to look at reference sets developed by other organisations. Existing reference sets may be useful for inspiration or they may be sufficient for either adoption or adaption. In either case exploring existing reference sets and evaluate whether these reference sets meet the requirements identified for this particular reference set should be considered:

- 6.2.1.1. Discover Existing Reference Sets
- 6.2.1.2. Evaluate Reference Sets

6.2.1.1. Discover Existing Reference Sets

In order to be able to adopt or adapt an existing reference set, you need to know what reference sets already exist. As there is no common library of SNOMED CT reference sets an authoring organization should apply alternative search strategies to explore what reference sets has already been developed.

Table 6.2.1.1-1: Overview of potential sources of reference sets that can be approached when seeking for existing reference sets

Potential source	Description
SNOMED International Browser	The online browser provide access to the International release of SNOMED CT and a range of national Editions, including reference sets defined within these Editions. The available reference sets should be represented as foundation metadata concepts and present as descendants of the concept 900000000000455006 | reference set (foundation metadata concept) |. SNOMED International browser provides access to the International Edition and some National Editions. The browser can be accessed here: http://browser.ihtsdotools.org/
National Release Centers	reference sets which are not included as national Extensions to SNOMED CT may also be acquired from the National Release Center. Each SNOMED CT Member has a National Release Center and information about these is provided by SNOMED International website, http://www.snomed.org/members/.
Personal networks/-Community of Practice	Explore your personal SNOMED CT network or the community of practice. The website www.snomedinaction.org provides access to information about existing SNOMED CT implementation initiatives. Some of these includes development and implementation of reference sets or subsets, so this is also a useful starting point for further exploration. Furthermore, attending the SNOMED Expo will also allow you to meet with persons directly involved with reference set development. You can find information about this event via SNOMED International website: http://www.snomed.org/participate/attend-ihtsdo-events
Web-based search	It may be useful to run a broad search on SNOMED CT reference sets or subsets using some of the great, commercial and non-topic specific search engines. This approach may result in identifying SNOMED CT subsets that use an alternative representation than the reference set format, or identifying clinical guidelines or recommendations that can help to specify the requirements for the content of the reference set. So, this approach is typically useful for getting inspiration to the content of the reference set to be developed.
Health IT vendors	It may very well be useful to explore what Terminology products and services are provided by Heath IT vendors. Providers of terminology servers and services may also have access to specific reference sets.

6.2.1.2. Evaluate Reference Sets

To determine whether a reference set is useful either to adopt or to build from, there is a range of parameters which can be assessed to determine whether a reference set is fit for the purpose or requirements that the developing organization have. The following diagram illustrates some of the central aspects that must be assessed when evaluating an existing reference set.

1. BASIC REFERENCE SET PARAMETERS

- The type of Reference Set
- The Number of members
- Extension content included
- Hierarchy coverage
- Common ancestors

2. CONTENT OF REFERENCE SETS

- False positives
- Level of specificity
- Oddness of any kind
- False negatives
- 'Edge cases'
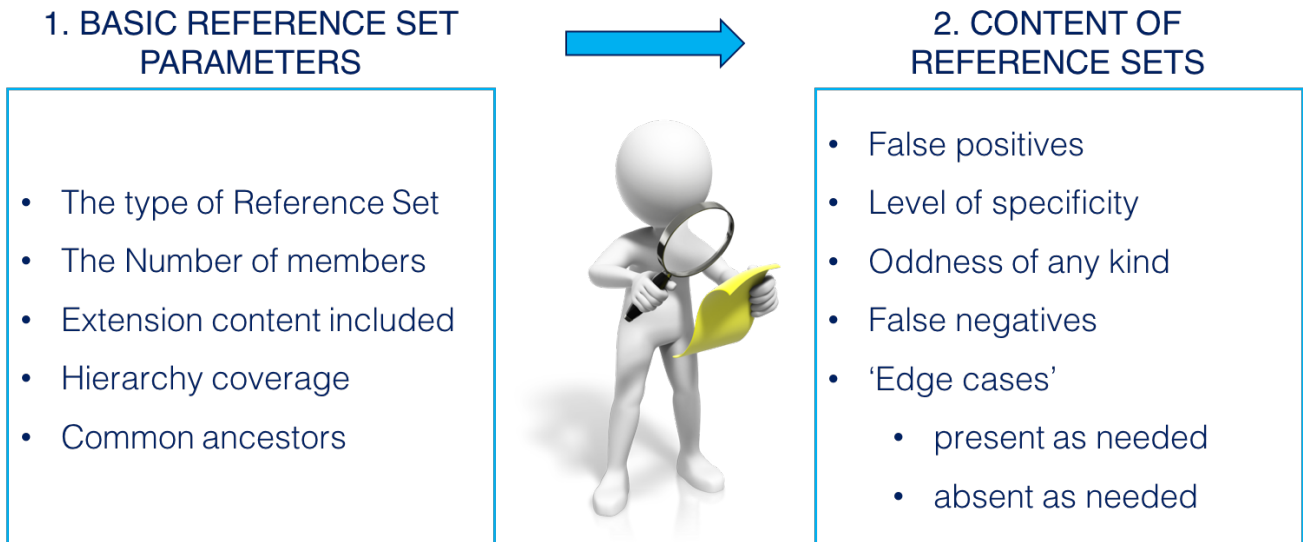 - present as needed
  - absent as needed

Figure 6.2.1.2-1: Exploring existing reference sets

Compare basic parameters

Firstly, the basic parameters of the existing reference set should be compared with the actual needs related to use of the new reference set. The following table provides an overview of some of the basic parameters, which should be assessed when starting to evaluate an existing reference set. Having assessed these basic parameters, it will be possible to determine whether an existing reference set is sufficient for further inspection, or whether it should be rejected as candidate source for a new reference set.

Table 6.2.1.2-1: Overview of basic reference set parameters relevant for evaluating reference sets

Parameter	Description
Type of reference set	It should be assessed whether the functionality of the existing reference set also meet the requirements for functionality that the seeking organization have.
The number of members	The numbers of members of the existing reference set should be compared to the required number of members in the new reference set. E.g. a reference set representing a very large subset of components may be rejected if the need is a subset to capture the overall categories (body sites) of information collected during a physical examination.
Extension content included	It should be assessed whether the existing reference set is dependent on any other Edition than the International Edition. And if so, the existing reference set can only be adopted if it belongs to an Extension on which the developing/seeking organization also depends upon.
Hierarchy coverage	It should be assessed whether the components referenced in the existing reference set represents concepts within hierarchies required for the new reference set. E.g. an existing reference set should be rejected if it references concepts within the Procedure hierarchy and the new reference set is to be used for recording evaluation results, i.e. concepts/descriptions of clinical findings concepts.

he more basic parameters it is also important to Inspect the specific content of the existing reference set to determine the appropriateness of this to fulfill the requirements that an organization have. The following table provides an overview of typical parameters to be inspected when evaluating the content of an existing reference set.

Table 6.2.1.2-2: Overview of typical aspects, which should be inspected when evaluating an existing reference set

Parameter	Description

False positives	Assess whether any of the components referenced in the existing reference set is inappropriate for the context of the new reference set and therefore should be excluded.
Level of specificity	Assess whether the components referenced in the existing reference set are at a sufficient level of specificity, e.g. assess whether concepts are too granular or too less granular to fulfill the requirements setup for the reference set.
"Oddness" of any type	Assess whether the coverage of the existing reference set meets immediate expectations. There might be outliers or content from other hierarchies than the primary domain for this reference set, which seems odd and therefore needs to be excluded.
False negatives	Inspect to seek out what is desirable to include, but is likely to be missing in the existing reference set. E.g. there might be requirements which reflect local needs or practices, which cannot be expected to be included in reference sets developed by another organization or in another context.
Edge cases	Seek out 'Edge cases' present as needed, absent as needed

anning and designing a reference set, the use of the reference set should be carefully analyzed. Specifically, it should be analyzed how the reference set is going to function with surrounding information models or software artefacts.

Bounded Reference sets

Bounded reference sets are designed to be used together with a specific information model. For example, if a reference set is developed as a value set for recording the smoking status in a specific software system, like shown in the example illustrated below.
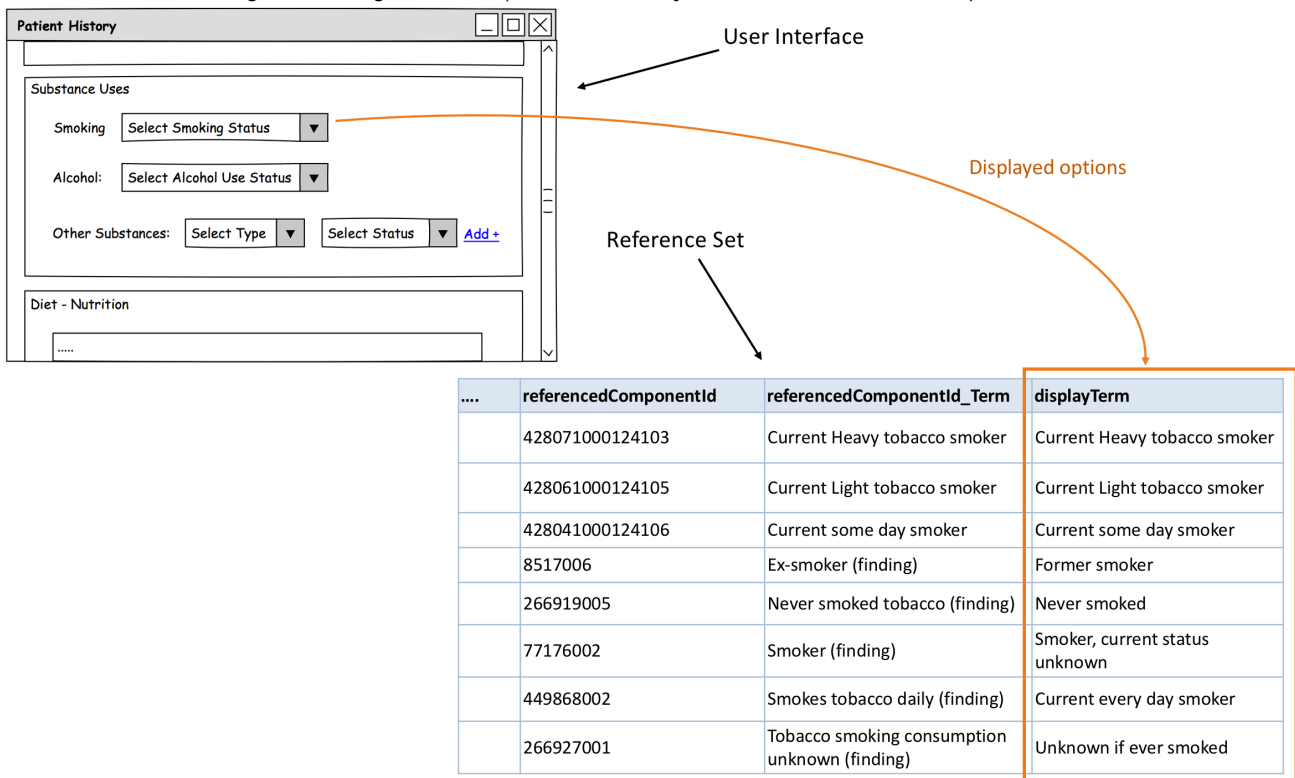
Figure 6.2.2-1: Some reference sets are designed to be bound for particular information models

When a reference set is bound to a specific information model, it is important to carefully consider how the binding affect the reference set members. So, for bounded reference sets it is important to clearly specify the relationship between the reference set and the associated information model to guide users, and to ensure correct interpretation of data, when data is subsequently retrieved for purposes such as display, analytics and communication.

Table 6.2.2-1: Questions to consider for bounded reference sets

Question	Examples

Does the binding modify the default context of the SNOMED CT concept used?	Negation or uncertainty • if a procedure concept is marked as planned • if a clinical finding is marked as absent Subject • if a clinical finding or a procedure is recorded in the context of a family member, or another subject, not being the subject of record For more information read the section about axis modification and about context representation 3.5. Safely representing the context of recorded codes.
What information model parts and related codes should be included when interpreting the meaning of data?	Different parts of the information model can be bound to SNOMED CT to express the meaning of the specific information model part. The Model as a whole, a group of data elements and each single data element can be bound to SNOMED CT. Dependent of which part of the model is bound to SNOMED CT, different methods are applied. • Model (model meaning binding) • Data group (Concept domain model meaning binding) • Data element (value set binding)

Unbounded Reference Sets

Unbounded reference sets mean that the reference set is designed to be applicable to multiple use cases, organisations or systems. An example is the SNOMED CT to ICD-10 map, which is released with the International Edition of SNOMED CT, and hence, available for use by any of the Members and affiliates. It may also be reference sets developed in a Member country to be used by all cardiovascular surgery departments in that country for reporting the procedures done and the procedure outcome. Such reference sets is distributed together with a National Edition of SNOMED CT.

For unbounded reference sets it is important for consistent and proper use to clearly specify the purpose and use of the reference set. Therefore, written specification and guidelines should be developed and distributed together with the reference set, to ensure that users of the reference set know how to implement and use the reference set.
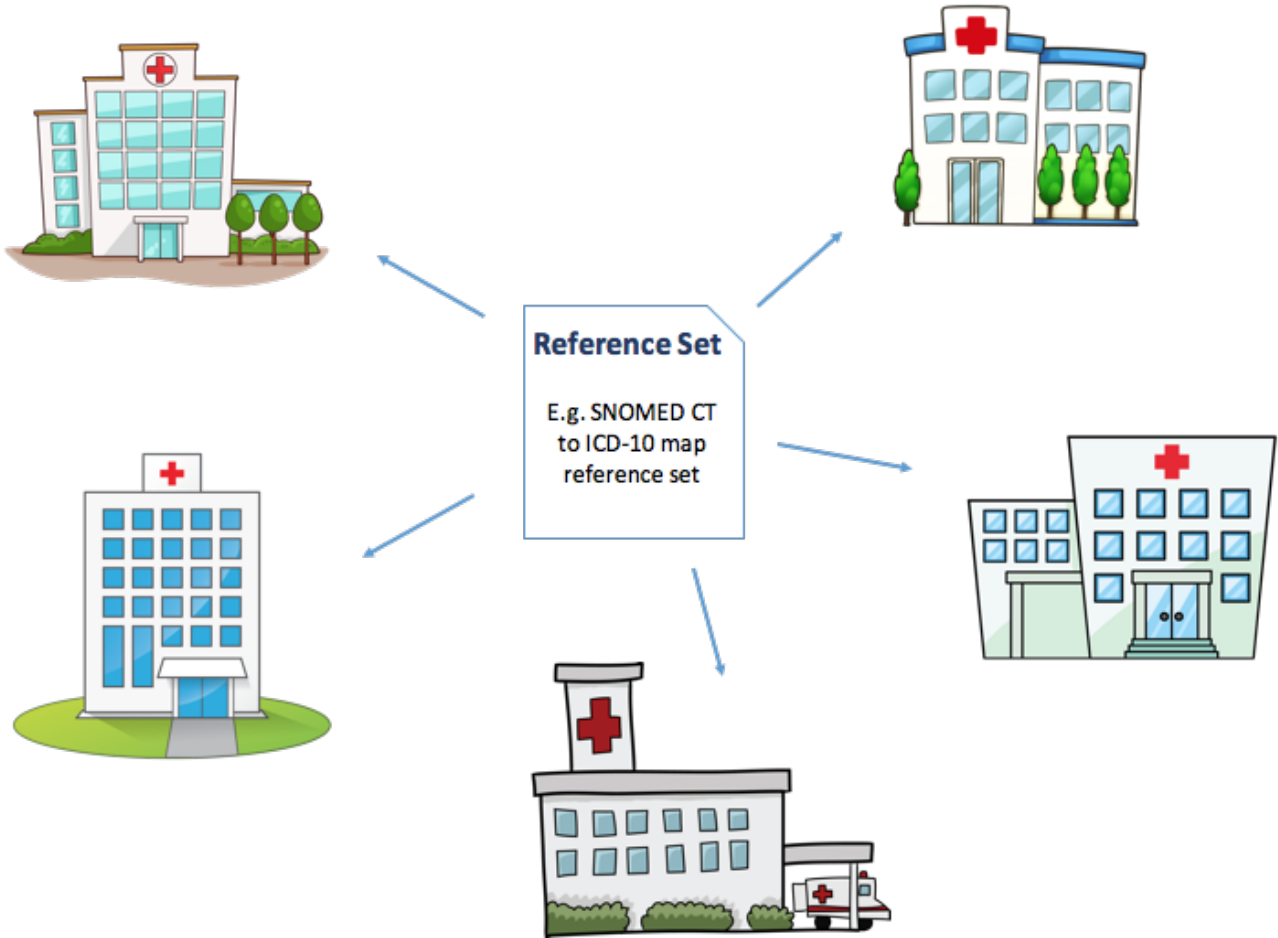
Figure 6.2.2-2: Unbounded reference sets are designed to be applicable to multiple use cases, organisations or systems

6.3. Development

In this part following topics related to the actual reference set development process is presented:

- 6.3.1. Reference Set Creation
- 6.3.2. Development Approaches
- 6.3.3. Development Methods
- 6.3.4. Review and Quality Assurance

6.3.1. Reference Set Creation

As part of developing a new reference set it should be created and identified as part of the Edition that it belongs to.

When creating a new reference set, the organization developing the reference set will need access to a namespace in order to generate SCTIds. Within their namespace, a moduleId concept (with an FSN and Preferred Term) should be added, placed in the 90000000000044 3000 |Module (core metadata concept)| subhierarchy within the metadata, for each authoring organization.

To specify any reference set it is necessary to identify the reference set by creating a concept and name the reference set by associated descriptions. The concepts should be a subtype of the relevant reference set type concept in the foundation metadata hierarchy. This could for example be as a descendant of the 446609009 |Simple type reference set (foundation metadata concept)| , or the 9000000000 00512005 |Query specification type reference set (foundation metadata concept)| . The latter is used for intensional reference set definitions. Following table shows the steps to follow when creating a new reference set.

Table 6.3.1-1: Tasks involved with creating a reference set

Step	Description
1. Define the reference set in the metadata hierarchy	1. Create a concept for the reference set 2. Add up to three Descriptions for the FSN, the Preferred Term and optionally the purpose, see 4.1.3. Naming Conventions for Reference Sets 3. Add an | is a | Relationship to link the reference set to the appropriate pattern
2. Define the reference set Attributes within the metadata hierarchy	Add new concepts for each of the reference set member attributes, if necessary. If the reference set attributes describing the pattern are adequate to describe the reference set's attributes, then these can be used instead. You may add new concepts for some of the attributes, and reuse existing concepts for other attributes, if you wish. Following steps should be taken for each attribute that you wish to create: 1. Add a concept for the attribute 2. Link it with an | is a | Relationship into the 900000000000457003 |Reference set attribute (foundation metadata concept)| hierarchy 3. Add Descriptions for each of the new attribute. If new attribute values need to be created these should also be added as SNOMED CT concepts and placed as subtypes of the concept 900000000000457003 |Reference set attribute (foundation metadata concept)|
3. Create the Descriptor for the reference set	If the reference set does not follow an existing reference set pattern, the additional attributes specific for this customized reference set should be included in the Descriptor reference set.
4. Add members to the reference set	Reference set members are added to the reference set, which includes specifying the attribute values for each reference set member.

For additional information, see: 13.1 How to create a new Reference Set using an existing pattern

6.3.2. Development Approaches

Developing reference sets can be done in different ways, depending on requirements, resources and skills of the organization who wish to develop a reference set. In some situation it may be the best approach to build the reference set without making use of or relying on any previous work for assistance. In other situations, existing reference sets may meet the requirements for another organization that wants to develop a reference set, and in these cases there are different approaches to utilize the existing work. In this section we introduce different general approaches to the creation of SNOMED CT reference sets, including:

- 6.3.2.1. Develop New Reference Set
- 6.3.2.2. Adopt an Existing Reference Set
- 6.3.2.3. Copy an Existing Reference Set
- 6.3.2.4. Adapt an Existing Reference Set

6.3.2.1. Develop New Reference Set

One way to create a reference set is to specify the content from SNOMED CT to be included in the reference set without looking at any existing SNOMED CT reference sets. This approach may be chosen, if no reference set is available that meets the requirements of the developing organization. It may also be because the requirements for the new reference set is so clearly defined or limited in scope that it is easier to create the reference set from scratch.

This could for example be, if a group of orthopedic surgeons are developing a subset of SNOMED CT procedure concepts to be included in a certain pick list of their local electronic health record. The group already know what options should be available in the pick list, so they choose to create their own reference set and add the required component references to this reference set, see figure below.
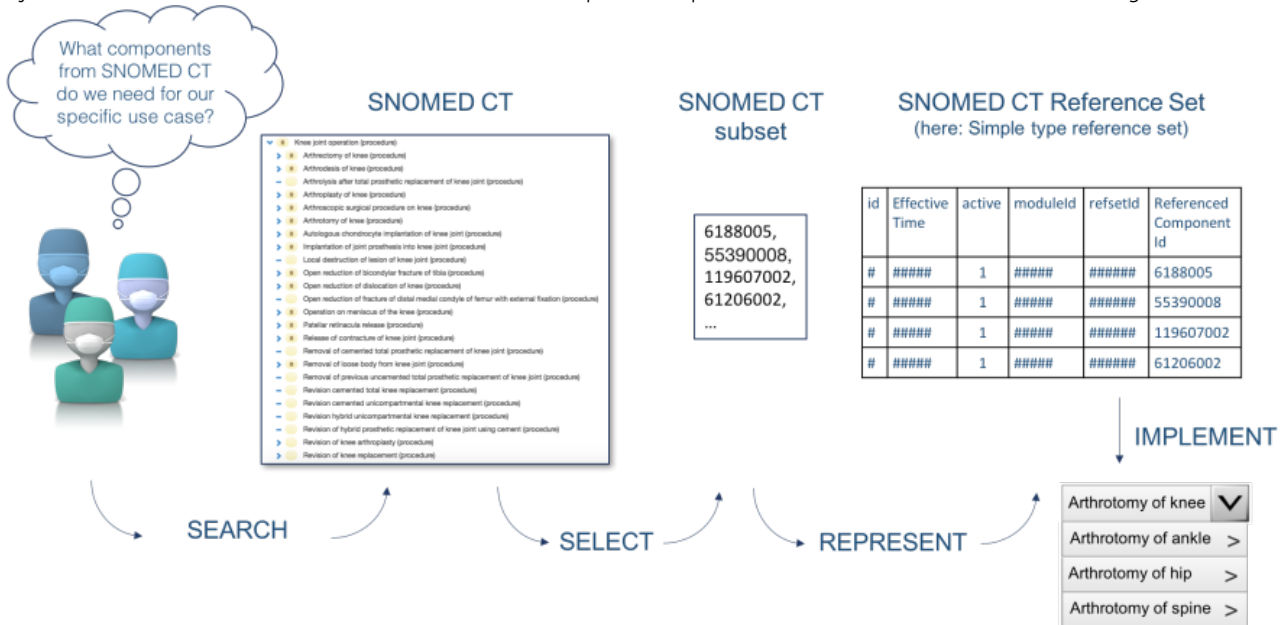
Figure 6.3.2.1-1: Creating a new reference set without looking at any existing SNOMED CT reference sets

Hence, you may choose to develop a new reference set if:

- No existing reference set meets you requirements
- The scope of the reference set is limited and clearly specified

Depending on the scope of content for the reference set there are different development methods that can be applied for selecting, or defining, reference set members.

Developing a new reference set also requires the developing organization to make changes to the reference set as necessary to meet evolving requirements.

6.3.2.2. Adopt an Existing Reference Set

As the experience of the SNOMED CT community grows, it becomes increasingly likely that there is some existing work that will provide a useful starting point, or that can be used to validate the work that you are doing. In this case, adopting an existing reference set may be a useful approach to take for acquiring a reference set.

Adopting an existing reference set is when the adopting organization/project use the existing reference set and future updates of that reference set as provided. This means, that the adopting organization commits to adopting any changes that are made to the source subset over time.

To be able to adopt and existing reference set it is a prerequisite that the existing reference set is part of a module that is included in the SNOMED CT Edition used by the adopting organization.

It is also a prerequisite that those maintaining the source subset publish changes in a predictable fashion, and the adopting project has processes in place to manage this adoption.

Figure 6.3.2.2-1: Adopt reference set

Hence, you may choose to adopt an existing reference set if:

- The reference set meets the requirements of your organization
- You are confident that the existing reference set will be well-maintained
- The reference set is part of a module that is included in the SNOMED CT Edition that you use.

Adopting an existing reference set is an attractive solution in situations where the source subset is already being used within the communicating community, and so avoids the need for mappings to be maintained, or other strategies developed to deal with differences.

The drawback is that if the adopting project requires changes to the contents of the subset it does not have any mechanism for requesting such changes or making them happen.

6.3.2.3. Copy an Existing Reference Set

In the situation where an existing reference set meets the requirements of an organization who wish to use a SNOMED CT reference set, it may be a solution to copy this reference set instead of adopting the reference set. Copying a reference set is a useful approach in the situation where the existing reference set is part of a Module which is not included in the SNOMED CT Edition that the copying organization use.

Copying an existing reference set means create a new reference set with members referencing the same components as the existing reference set.

It is important to have a clear strategy for maintenance of a copied reference set. When an organization choose to copy an existing reference set they need to make changes to the reference set as necessary to meet their evolving requirements. Whether it is the copying organization or the authoring organization who is responsible for adding or inactivating content to the existing reference set depends on the agreement between the involved parties. It also depends on whether the authors of the existing reference set have established a process, which deals with requests for changes. In either situation the copying organization will need to apply the changes made to any new version of the original reference set to the copied reference set.
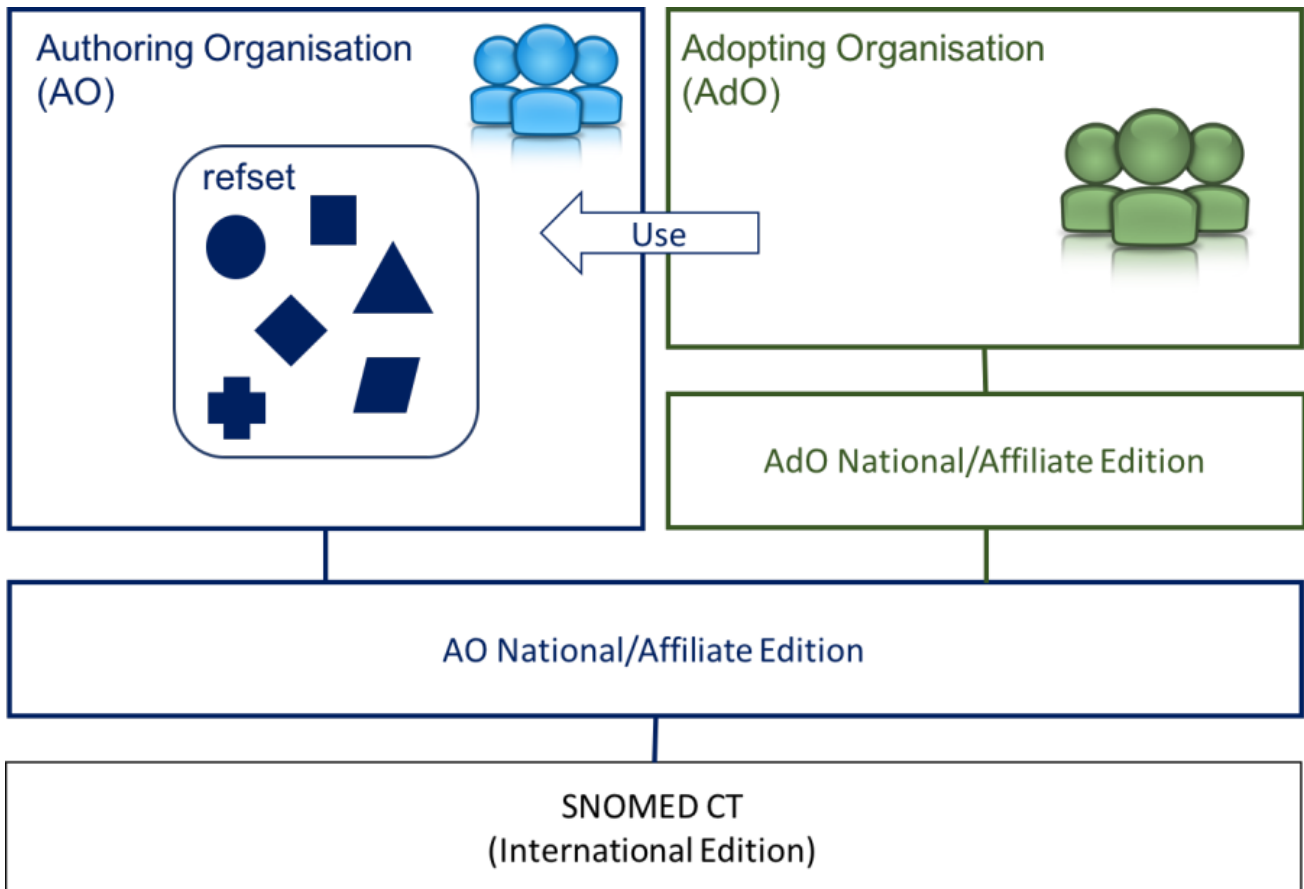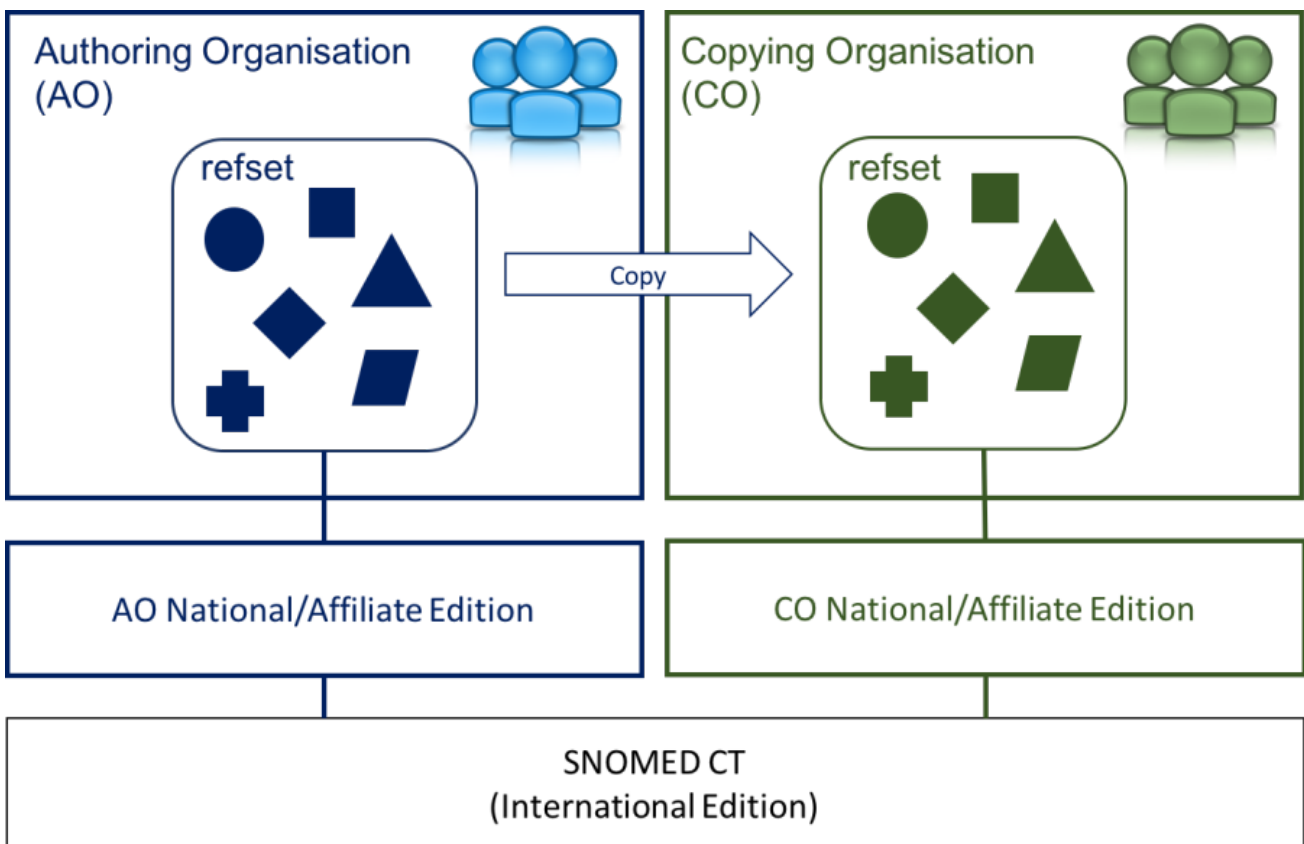

Figure 6.3.2.3-2: Copy reference set

Hence, you may choose to copy an existing reference set if:

- The reference set meets the requirements of your organization
- You are NOT confident that the existing reference set will be well-maintained
- The existing reference set is NOT part of a module that is included in the SNOMED CT Edition that you use.

6.3.2.4. Adapt an Existing Reference Set

Another approach to develop a reference set, is to adapt an existing reference set. This is where and existing reference set is used as a source of inspiration for the developing organization. It may also be, that the existing reference set almost meet the requirements of the adapting organization, but some modifications are needed. The adapting organization may wish to add more content to the existing reference set, or there may be content in the existing reference set which should be excluded from the new reference set.

Like copying reference sets, adapting a reference set will include the creation of a new reference set which is authored and maintained by the adapting organization. The members of the new reference set may then reference components that are also referenced in the existing reference set, they may exclude references that are in the existing reference set, or they may add other component references to meet the requirements of the adapting organization.
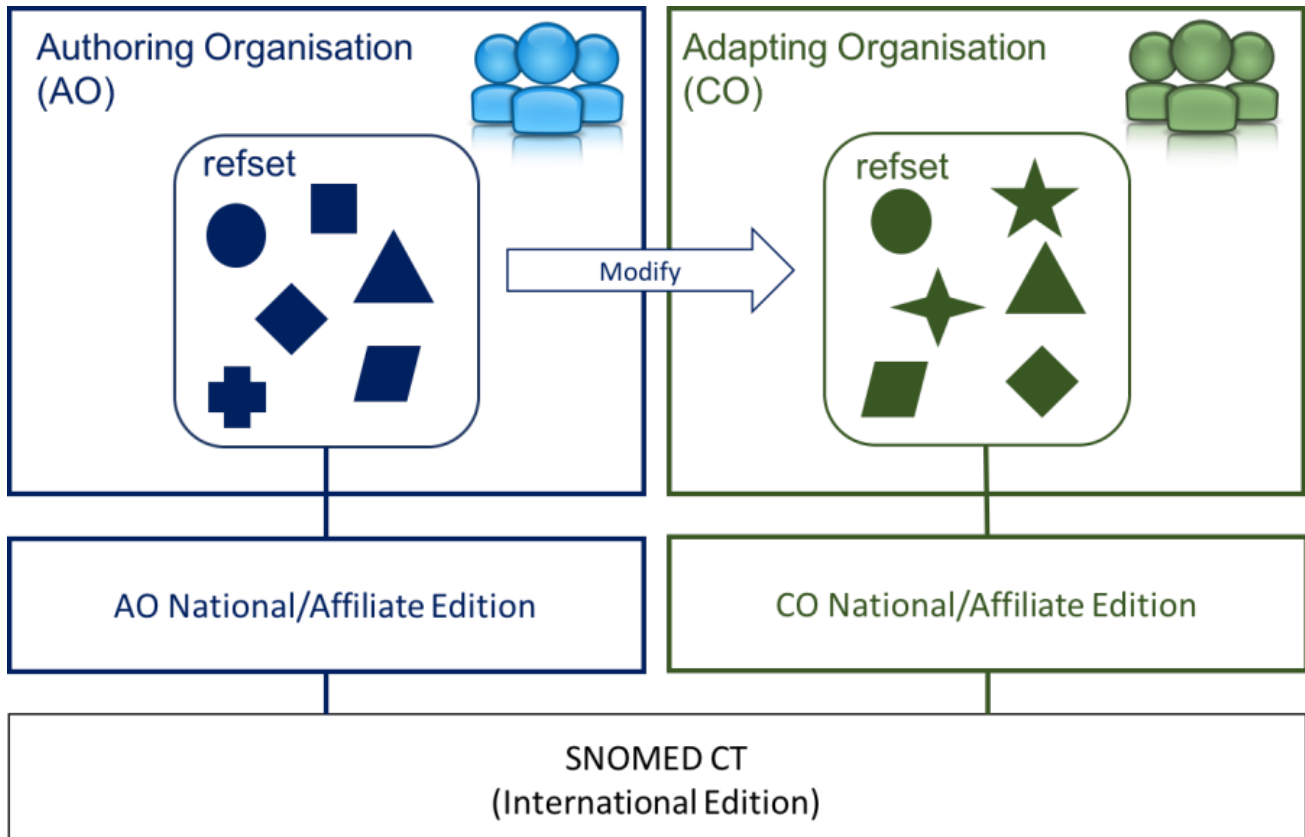

Figure 6.3.2.4-1: Adapt reference set

Hence, you may choose to adapt an existing reference set if:
- The reference set almost meets the requirements of your organization, but you wish to modify it to fully fulfill your requirements.

6.3.3. Development Methods

When developing simple reference sets, there are different methods which can be applied to determine what concepts to be included in the subset. The two overall approaches to defining the reference set members are manual enumeration and intensional definition.
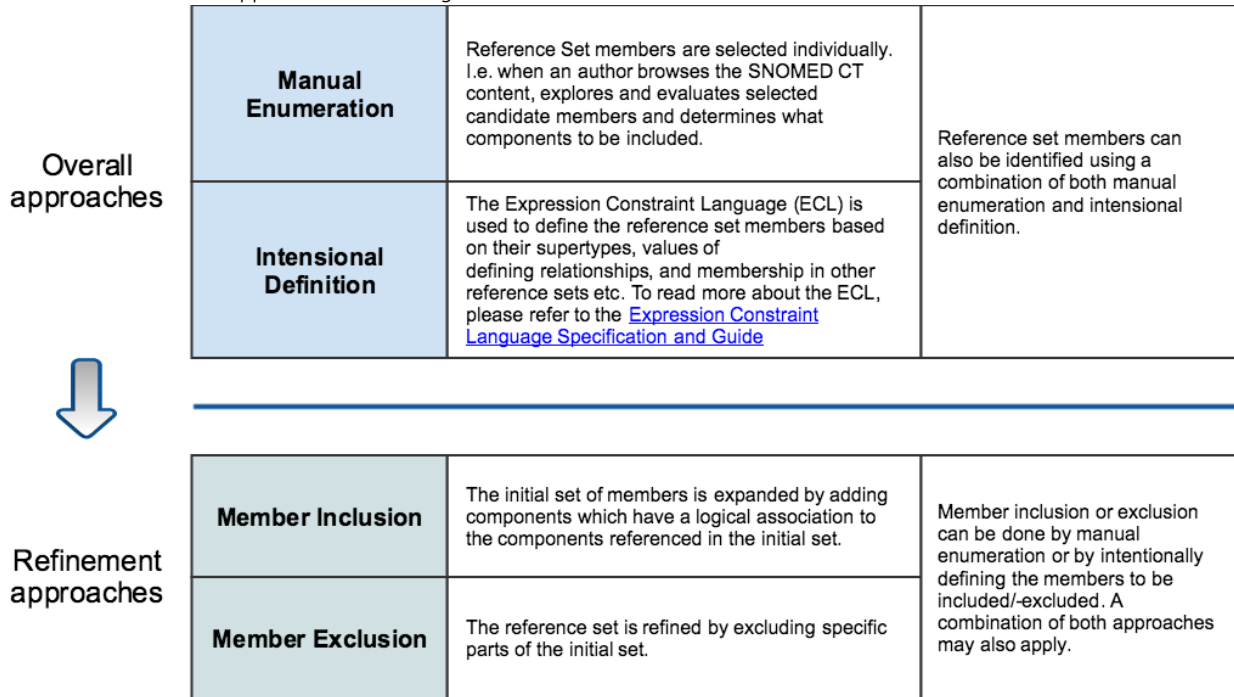
Figure 6.3.3-1: Reference set development approaches

Requirements for Manual Enumeration

Identifying reference set members by manual enumeration is useful when the number of members required is limited or when migrating from a well-specified domain. The source information may come from a variety of sources, such as:

- Subsets of concepts from other code systems
- Legacy codes (when migrating from a system that do not use SNOMED CT)
- Clinical guidelines and other knowledge resources

Manual enumeration is also useful when the reference set should be stable over time, i.e when addition of concepts, in subsequent releases of SNOMED CT, is not required to be reflected in the reference set.

Requirements for Intensional Definition

Identifying reference set members intensionally is useful for creating subset of concepts who share a set of common characteristics, such as:

- Being a descendant of the same focus concept
- Sharing one or more attribute relationships

Intensional definition of reference set members is also useful when changes to concepts, in subsequent releases of SNOMED CT, is required to be reflected in the reference set.

6.3.4. Review and Quality Assurance

Ensuring the quality of a reference set is important for the successful use of SNOMED CT. It is crucial that the content of the reference set references the concepts which represent the meaning of the data it is to be linked to. In addition, it is important that the human interpretation of the concepts referred to aligns with the logical definition provided by SNOMED CT. Even when using reference sets for less patient safety critical purposes, the quality of the reference set needs to be sufficient to be trusted to serve its purpose.

Stages of Quality Assurance

Reviewing a reference set is important throughout the reference set development process, however at least three types of validation should be emphasized, see illustration below.

- Design review: The overall objective with this review is to verify whether the reference set design meets the requirements.
- Content review: The overall objective with this review is to verify whether the selected reference set members are sufficient for the context where the reference set is to be used.
- Test: The overall objective with testing the reference set is to validate the reference set in the context where it is to be used. Testing is done to assure that the reference set meets the needs of the involved stakeholders.
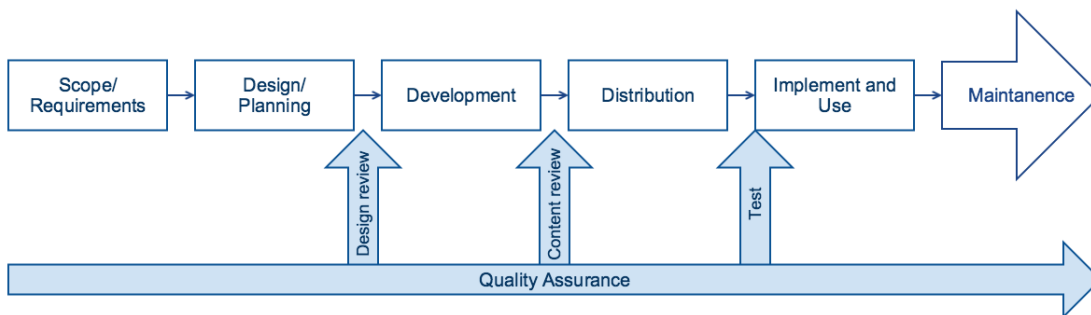

Figure 6.3.4-1: Quality assurance stages

Review

Reviewing a reference set before testing it in the context of its usage is important in order to correct the most distinctive errors, such as content gaps or disagreements about selection of concepts. Additionally, reviewing a reference set is important, because it is much easier to make major adjustments as early as possible. In addition, the adaptability of the reference set increases the higher the quality of the reference set, according to the perspective of the subset users. The level of quality assurance needed will of course be dependent on the practical use case of the reference set.

The checks that should be done as part of the review process include:

1. Check that all concept ids are current (that there is no references to inactive concepts)
2. Check that all concepts included in the reference set are appropriate to the use case
3. Check that the terms used to describe these concept ids are appropriate
4. Check that there are no gaps in content. I.e. concepts that should be included in the reference set that aren't
5. Check that there is clear context and meaning, which align with the surrounding models that the reference set is to be used with

Skills and Roles

The people involved with reviewing the reference set should be by a combination of domain experts and SNOMED CT experts.

- Domain experts: There may be different types of domain experts, but at least two types of domain experts can be specified :
  - Clinical domain experts: Persons who have knowledge about the use of the reference set and knowledge about the context where the subset is going to function. Often, clinical domain experts will be clinical users who have the sufficient level of clinical knowledge, to evaluate whether the scope and the content of the reference set is sufficient.
  - Technical domain experts: Persons who have knowledge about the technical environment where the reference set is going to be implemented. technical experts should have knowledge about the various models to which this reference set should be bound. Often, these experts will have a technical background, for example software architects, programmers, IT-managers. These experts are especially important at an early stage of the reference set development process, to review and validate the design of the reference set in relation to the requirements set out by the technical infrastructure.
- SNOMED CT Experts: Persons who have the sufficient knowledge about SNOMED CT to review the concepts selected for a specific purpose. These persons should be able to assess whether the definition of the selected reference set members suite the

domain of the reference set. For example, if the reference set represents a problem list, then the included members should reference clinical finding concepts. This group is also serve an additional, supportive role in terms of guiding the domain experts in reaching correct interpretation of the reference set structure and content.

Review approaches

SNOMED International does not recommend ONE specific approach for review of reference sets, but blinded approaches are typically a feasible approach to achieve high quality reference set. Alternatively, or additionally, a combination of review methods can be recommendable to ensure a reliable and usable review process. Ideally, reference sets should be reviewed iteratively until the reviewers are satisfied. Examples of the approaches to take are illustrated in the following table.

Reference set review approaches.

Approach	Description
Single author single reviewer	In this approach one person author the reference set, i.e. determines the content to be included, while another person reviews the selected reference set members. The reviewer may also add comments against the proposal and suggest alternative content. In a really small team, the author and reviewer may be the same person. However this is not ideal, and it is highly recommended to include two or more people in the review process.
Cross-review	Using this approach the reference set development work is divided evenly between two or more authors of the reference set. 1. The authors divide the material between, so they are each responsible for the initial development of a specific part of the reference set. 2. The authors then swap their material and they each review each other's material. This approach can be very efficient if time is short, because you can spread both the development and the review load between more than one person – however, reviewing someone else's material may not always as effective (in terms of quality) as the dual blinded approach.
Dual blind review with an adjudicator	This approach is useful in a slightly bigger team than required for the single mapper single reviewer approach. This approach includes following steps: 1. Two authors develop a draft reference set based on exactly the same material 2. Their reference sets are then compared, to identify any differences. If any differences are found, then 3. The adjudicator compares the reference set differences and decides which is appropriate. In some cases, the adjudicator may ask the authors to provide their reasoning for their choice of reference set members, and this may be used to help make the decision. This approach can produce higher quality reference sets, because each author is independently cross-checking each other's material without being biased by the decisions of the other author. While this approach takes longer during the development phase (because both authors need to develop a draft of the whole reference set), the review phase can be a lot quicker (because the adjudicator only needs to check the parts that the authors disagreed on).
Workshop	Validation workshops is workshops dedicated to review and validate the design and/or content of a reference set. In these workshops the content or uncertainties are discussed in details, or test-persons are asked to prioritize and assess specific subset members etc. The participants may have had a chance to review the reference set individually prior to the workshop to prepare questions and comments for discussion. The number of people in the workshop and their roles should be considered and selected dependent of the format and the scope of a specific workshop. This approach is time consuming which should be acknowledged already in the planning stage. However, this approach may also be rewarding. Workshops often give rise to detailed discussions or unplanned discussions of relevant issues, but at the sametime workshops provide an opportunity of increased ownership and participation among the participants, which may have a positive effect on the adoption of the reference set. It is recommended to plan these workshops in detail and to include a set of workshops. The number of workshops necessary depends on the size of the reference set, and how the feedback sharing is conducted.

Test

In addition to the abstract assurance of the subset in isolation, there may be a requirement for a level of testing to be undertaken in healthcare systems, and tested under the exact circumstances of intended use. This may be undertaken by releasing a technology preview targeted at specific bodies for feedback. An impact assessment and, most importantly, safety testing will need to take place upon deployment of the subset into systems, particularly where significant changes have taken place.

6.4. Distribution

Reference Sets can be distributed as part of the International Release, as part of a National Edition or as part of an Affiliate Edition, and the way the reference set is distributed will typically depend on how the Edition is distributed. However, in some situations, reference sets may be distributed independently.

Distributing reference sets usually involves performing any additional quality checks required to ensure that the data can be exported correctly, and then exporting the reference set from the terminology management tool. Finally, the reference set is distributed to the users, for example as release files accessed from an online library using a terminology server API.

Distribution Format

The standard format for distributing SNOMED CT reference sets are the reference set format, as described in 4.2. Common Reference Set Format. For example, an extensional definition of a subsets of components is represented as a simple reference set, and the standard format for distributing intensional definitions of concept subsets is as a query specification reference set. As part of the SNOMED CT release format, reference sets support unique identification, versioning and recognition of dependencies.

Using the reference set format for distributing SNOMED CT derivatives is beneficial from the view of maintenance, because the versioning attributes is identical to the versioning attributes used for SNOMED CT components. This means that the checks that need to be done as part of a regular release cycle is easier to manage compared to having a range of different versioning mechanisms to adjust to.

Other distribution formats may be used where necessary. For example to comply with requirements for representation of value sets including codes from other code systems. However, care should be taken to ensure that distributed subsets are uniquely identified and versioned. Subsets need to be accessed, selected and where appropriate bound to information models using standard approaches.

Release Cycle

Some subsets will be distributed biannually as part of the International Release. However, some others may not require this release schedule. Release processes must include a period of time dedicated to subset manipulation, following the 'freezing' of SNOMED CT itself. This is the period where all the changes has been made to the core components of SNOMED CT. Timing of release and distribution of specific subsets will be dictated by individual use case-specific requirements. These should be identified and documented initially as part of the subset development process. These may change over time based on user feedback.

6.5. Implement and Use

When a reference set has been developed and tested it should be prepared for use by implementing it in the environment where it is going to function. I.e. the implementation process and the tasks involved will vary dependent on the setting.

Often, implementation of a reference set will mean integration with one or more software artefacts, to enable data capture, represented by appropriate reference set members.

An important part of this process is proper binding between the information model and the reference set to ensure effective integration with data entry and storage functions. Many situations will also involve implementation of two or more interdependent reference sets, which require their dependencies to be considered to reach the most optimal implementation.

In this guide we highlight two aspects that are important when implementing SNOMED CT reference sets, i.e. supporting implementation with proper implementation guidance and considerations related to optimizing the reference set for a particular implementation.

- 6.5.1. Implementation Guidance
- 6.5.2. Optimization for Implementation

6.5.1. Implementation Guidance

The tasks involved with implementing a reference set will depend on the situation and the technical setting where the reference set is going to function. To support successful implementation, and to ensure that the reference set is used as intended it is important with precise and clear instructions for the people responsible for implementing the reference set, i.e. the people preparing the reference set for routine use within a particular setting.

Implementation guidance should, as a minimum, include instructions on:

- The overall purpose and intended use of the reference set
- The reference set design, i.e. attributes and data types
- Dependencies to terminology artifacts, such as
    - other reference sets
    - content in the international edition of SNOMED CT
    - content in SNOMED CT extensions
- Dependencies to software artifacts, such as
    - information model structures, i.e.
        - data entry template structures
        - storage model structures
        - message structures
- Maintenance
    - release cycles
    - change management
    - request for change

Additionally, it is valuable to include instructions on how to test the implementation, i.e. exemplar test cases with information on expected outcome given a specific action.

6.5.2. Optimization for Implementation

As part of an implementation process it is relevant to consider any optimization that could be done to meet requirements for implementation or result in a more effective and/or efficient use at runtime. Because, even though reference sets are distributed in separate files and represented according to the reference set file format, it may be useful to transform these reference sets files into another structure that is more appropriate for the specific use case.
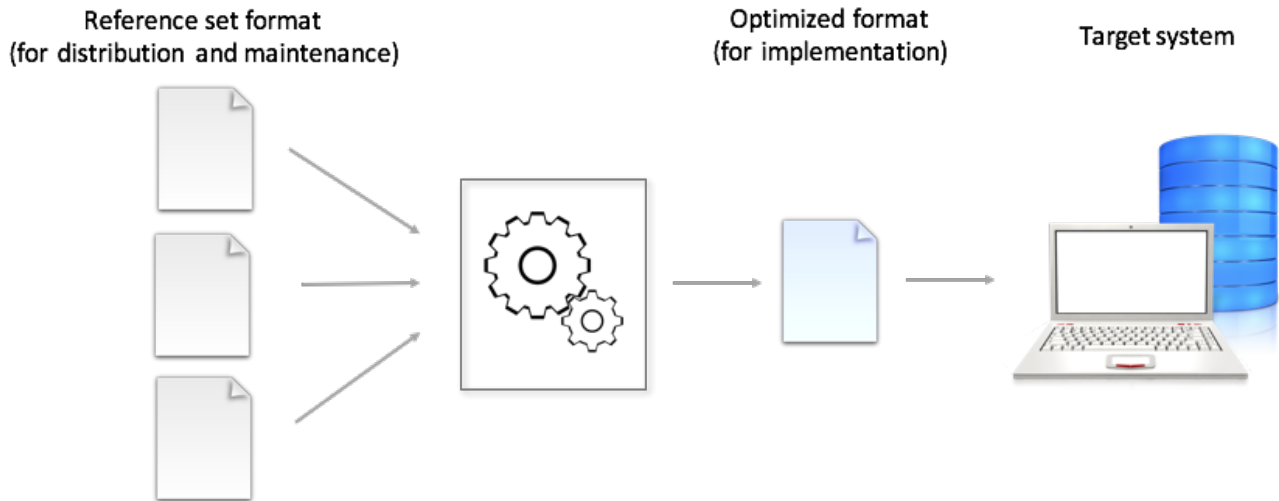

Figure 6.5.2-1: Prior to implementation a reference set may be transformed into one or more formats, which are optimized for the particular use of the reference set

Optimizations that could be considered include:

- Controlled redundancy. For some matters it may be useful to introduce some level of redundancy, e.g. by joining reference sets or release files together so that the data is more efficient to access at runtime
  - e.g. combining a simple reference set of concepts with the description file, so that the term preferred for display can be retrieved directly from the subset file.
  - e.g. combining the description file with a language refset, so that the term is in the same row as the acceptability of the description, and the concept id that it describes.

- Terminology reduction (constrain). For the purpose of supporting effective use and ensuring that only relevant components are accessed during runtime it may be useful to filter the components down to only those that are relevant for a specific use case.
  - One way of doing this is by creating simple reference sets of the components that are relevant in a given situation, for example:
    - A subset of relevant concepts.
    - A subset of descriptions acceptable or preferred for the set of concepts.
    - A subset of relationships that are important to know about for that specific subset of concepts and for the given use case. What relationships to include will be very dependent on the use case. In some situations, the relationship subset may be irrelevant, and in other situations it may be important to include all defining relationships and all transitive 'is a' relationships of the subset members.

- Filtering to a snapshot view of the subset. As for any other SNOMED CT component changes to reference set members are likely to occur, and as for SNOMED CT components, reference set members can be updated or inactivated. Therefore, it may be useful to create a snapshot view of the reference set to be used for implementation, so that only the latest version of all reference set members can be accessed during runtime. Alternatively, it may be useful to create a view which only includes the most recent version of all active reference set members. This will depend on whether the inactive members should be accessible for viewing or not.

6.6. Manage and Maintain Reference Sets

With time, the set of concepts, or other components, referenced in a reference set will need to change, either because of changes in SNOMED CT, changes in clinical knowledge, or changes dictated by errors or omissions in the original reference set. The original reference set will have been given an identifier that is used to reference the reference set in models or other entities that use that particular reference set. When a change is required that maintains the original intent but updates the content, the change is considered a new version of the reference set. Versioning maintains the reference set identifier and additional reference set attributes.

Following topics will be presented in relation to reference set maintenance:

- 6.6.1. Reference Set Management Considerations
- 6.6.2. Managing Reference Set Changes
- 6.6.3. Request for Change
- 6.6.4. Tooling

6.6.1. Reference Set Management Considerations

When developing or evaluating a change management process it is important to take different factors into account. Some of these factors are introduced in the table below.

Table 6.6.1-1: Change management considerations

Topic	Description
Request for change	Where will the need for change come from?
	This could be any of the stakeholders involved in the capture or use of the information that is expressed using the subset.
	How will change requests be expressed and submitted?
	These may be submitted directly by users, or may be collated and edited by suppliers. The maintainers of the subset may be proactive in looking for improvements, or may wait for requests for change to be submitted.
	What lead time is acceptable for the processing of a change request?
	What should users of the subset do in the period between recognizing the need for a change, and that need being met by a new release of the subset?
	Options may include using free text, creating local codes, or waiting for the revision to be available.
Revision cycle	Will there be a predictable revision cycle with regular releases, or will changes be made on an as-needed basis?
Resources	What editorial and technical resources are needed?

6.6.2. Managing Reference Set Changes

Maintaining reference sets starts as soon as the reference set is specified. The maintenance process involves ensuring that the content of the reference set refers to active concepts in the release upon which the reference set depends. It also involves adding new content to the reference set, for example when additions to the dependee Release fall within the scope of the reference set specification. The alignment of a reference set with the International Release will be undertaken during the international release process, at which time a change report will be created for circulation, providing details of the changes made, in a format that can be circulated virtually to subject matter experts.

The tasks related to reference set changes will typically involve substituting a prior version of a reference set with the more recent version. For example, if a reference set represents a value set which is used in a data entry interface, the reference set may simply be replaced by the updated version of that reference set. On the other hand, if the reference set is of a more complex nature, for example, if it represents a map between SNOMED CT and a classification, and is used for quality monitoring, research, reimbursement etc., it might be necessary to go into more detail about the implications of the change. Regardless of the type of reference set change (addition, inactivation or change) it will always have implications.

Addition of Reference Sets

If a reference set is added data entry protocols should be updated to link to this reference set where this is found appropriate. For example, if a reference set has been developed to function as a value set for cardiovascular diseases, data entry protocols should link to the reference set in the data entry protocols used to capture cardiovascular diseases. The nature of the binding depends on the underlying information model of the protocol. Additionally, it is important to update the storage models appropriately. See, Starter Guide Chapter 3. Using SNOMED CT in Clinical Information to learn about this topic. It may also be necessary to develop queries to generate reports or views based on the new reference set data. Alternatively, an added reference set may include rules to enable decision support, which must be integrated. It may also be necessary to create or update links between reference set members and communication protocols.

Inactivation of Reference Set

A reference set may be inactivated in situations where an organization is no longer able or prepared to maintain it.
Inactivation of an entire reference set is done by inactivating the concept that identifies the reference set, and inactivating the relationship used to place the reference set in the subtype hierarchy.
Reference sets that are being actively used should not be inactivated without giving prior notice to the users of the reference set, to allow remedial steps to be taken.
Any function using a reference set that is no longer maintained will need updating. For example, if a reference set is inactivated, which is used for data entry, it is important to determine how the inactivation is managed in the system. Thus, it should be clarified whether a new reference set should be defined or whether the inactivated reference set should be added in a local extension. It may also be that a replacement reference set has already been defined.

Changes to Reference Set Members

The types of changes that can occur in a reference set include:

- Changes to reference set members
- Addition of reference set members
- Inactivation of reference set members

Additional information: TSG Guide Section 13.2 How to add, change or remove members of an existing reference set

The effect of changes to reference set members depends on the type of reference set and the way the reference set is used. Generally, it is important to assess the extent to which the reference set changes affect the situation in which the reference set is used, such as:

- Data entry. Changes to reference set members may require changing existing data entry protocols.
- Data storage. Changes to reference set members may require migrating or processing pre-existing data in a particular way to ensure consistency.
- Data retrieval. Changes to reference set members may require revising existing queries to take account of the changes.
- Knowledge linkage. Changes to reference set members may require updating bindings to knowledge resources or decision support rules.
- Communications. Changes to reference set members may require updating communication specifications, and in particular managing issues from cross-version communications.

Determine changes

One way of determining the changes to a reference set is by creating and comparing the following two views of the reference set:

- SNAPSHOT view of the previous reference set release. This view contains one version of every member released up to the time of the snapshot. The version of each member contained in a snapshot is the most recent version of that member at the time of the snapshot.

- DELTA view of the new reference set release. This view contains only reference set member versions created since the previous reference set release. Each reference set member version in the DELTA view represents either a new reference set member or a change to an existing member.

These two views make it possible to automatically identify whether a reference set member has been created, changed, reactivated or inactivated. The table below illustrates how to interpret the 'active' attribute, when comparing the delta view of the new release of the reference set with the snapshot view of the previous release of the reference set.

		Value of active column in the new release of the reference set (DELTA view)		
		0	1	NOT PRESENT
Value of active column in previous release of the reference set (SNAPSHOT view)	0	Inactive member changed (not significant)	Member REACTIVATED	NO CHANGE
	1	Member INACTIVATED	CHANGED member	NO CHANGE
	NOT PRESENT	n/a	NEW member	n/a

6.6.3. Request for Change

As clinical practice evolves and adapts to meet the needs of healthcare advances, SNOMED CT content is updated to address these changes. Reference sets may also need to be updated to reflect content changes and to enable clinicians to adequately represent new and emerging techniques and practices. It is therefore important to establish processes for end users and other stakeholders to submit requests for changes. Related to reference sets the types of changes that may be required include:

- Addition of new reference set members
- Removal of reference set members
- Changes to existing reference set members

Types of Reference Set Changes

Addition of New Reference Set Members

When new members are requested for inclusion in a reference set, it should be determined whether any existing components in SNOMED CT are sufficient, or whether it requires a new component to be added - either in the International Edition or as part of an Extension.

Removal of Reference Set Members (Inactivation)

To remove a referenced component from a reference set, the relevant reference set member is inactivated. This is done by adding a new row to the reference set in which the value of the active column is set to false. The SNOMED CT versioning mechanism tracks the full history of additions, changes and inactivations because each row in the reference set contains a effective time column indicating when the change became effective. This is the same versioning mechanism with is used to track addition, inactivation and modification of SNOMED CT components.

Changes to Existing Reference Set Members

Changes to reference set members can vary, dependent on the refset type and the situation. Examples of changes include:

- Changing the order assigned to a member of an ordered reference set
- Changing the annotation associated to a member of an annotation reference set
- Changing a description from being acceptable to preferred in a language reference set
- Updating the query for an intensional subset definition represented by a member of a query specification reference set

Approaches

The type of service implemented to support change requests will depend on the amount of reference sets, the number of users, the organisation etc.. Smaller organisations with relatively a few end-users may apply a simple email service for proposing any changes to the refset. Other institutions may implement dedicated systems that track each change request and support proper management of those.

6.6.4. Tooling

Authoring and managing reference sets require some level of tooling support. Broadly speaking, two types of tooling support exists:

1. Combination of browsing functionality and simple spreadsheets or database system
2. Dedicated reference set management tools with extended features

The extend of sophistication of the tool required depends on various factors, and the combination of those factors:

- Type and size of reference set
    - Simple reference sets with few members
    - Large reference sets of various types
- Use of reference set
    - Local vs. distributed
    - Dependencies (Editions, versions)
- Maintenance of reference set
    - Cycle
    - People

Functionalities

Relevant functionalities for reference set creation and management include:

- Creation of both extensional and intensional reference sets
    - Services for browsing and displaying SNOMED CT content for manually selecting reference set members
    - Services for manual and intensional expansion of reference set (inclusion of reference set members)
    - Services for manual and intensional exclusion of reference set members
    - Create reference sets according to a specific reference set pattern
    - Create customized reference sets
- Version management of created reference sets
    - Computation of the effect of a related SNOMED CT edition and version on a reference set
- Workflow support
    - Mechanism for approval of reference set before publishing
    - Facility for external feedbackand request for change
- Searching existing reference sets, as well as entries about reference sets that may be available elsewhere
- Distribution services

Considerations Related to using Spreadsheets to Manage Reference Sets

Generally, it is not recommended to use a spreadsheet approach for managing reference sets, because there is a dependency between the reference set and the core content of SNOMED CT. The SNOMED CT release files are important for interpreting the components referenced by the reference set, for example:

- Description file: Access to the descriptions are required to display the terms of the components referenced by the reference set.
- Concept file: Access to the concept file is required to assess whether the components referenced by the reference set are active or inactive

However, in some situations spreadsheets (supported by a browsing facility) may be feasible, for example,

- If the reference set only consists of a few members
- If the reference set is only used locally
- If only one person is responsible for managing the reference set

If a spreadsheet approach is chosen it is recommended to create a human-readable version of that reference set to support interpretation of the identifiers used in the reference set.

Appendix 1: Overview of Reference Set Types

Content Reference Sets

Table 1: Overview of reference set types and example use cases, as described in the international Release of SNOMED CT

Reference Set Type	Description	Example Use Cases
Simple reference set	Allows a set of components to be specified for inclusion or exclusion for a specified purpose	Constrain values available in clinical data entry templates. (E.g. define pick lists, constrain searches etc.) Specify values accepted for communication purposes in specific elements in a communication message
Ordered reference set	Allows a collection of components to be defined with a specified given a priority ordering	Alternative navigation hierarchies. Specify a preferred order of concepts/descriptions in pick lists
Attribute value reference set	Allows a value from a specified range to be associated with a component	This reference set type can be used for many different purposes that are related to both content and technical use cases. E.g. to specify why a concept or description has been inactivated
Simple map reference set	Allows representation of simple maps between SNOMED CT concepts and values in other code systems.	Appropriate where there is a close "one-to-one" mapping between SNOMED CT concepts and coded values in another code system.
Complex and Extended Map reference sets	Allows representation of simple complex maps between SNOMED CT concepts and values in other code systems	Enables representation of maps where each SNOMED CT concept may map to one or more codes in a target scheme, or where the correlation of each map should be specified
Language reference set	Supports the representation of language and dialects preferences for the use of particular descriptions	Used to specify the acceptable and preferred terms for use within a particular country or region. Can also be used to represent preferences for use of descriptions in a more specific context such as a clinical specialty, organization or department
Query specification reference set	Allows a serialized query to represent the membership of a subset of SNOMED CT components	Used to represent Intentional definitions of reference sets. Constrain values available in clinical data entry templates, as part of a terminology binding
Annotation reference set	Allows text strings to be associated with components for any specified purpose	E.g. linking a SNOMED CT component to a url, e.g. for linking to a clinical guideline
Association reference set	Represents a set of unordered associations of a particular type between components	Used to associate inactive concepts with active concepts that can serve as potential replacements for the inactivated concepts

Reference Sets for Technical Use

Table 1: Overview of reference set types and example use cases, as described in the international Release of SNOMED CT

Reference Set Type	Description	Example Use Cases
Module dependency reference set	Represents dependencies between different SNOMED CT release modules	The Module Dependency reference set is used to ensure that all dependencies are satisfied when importing data
Description format reference set	Specifies the text format and maximum length of each supported description type	Specify new, localized description formats to be used in an Extension
Reference set descriptor reference set	Represents the format of all reference sets included in a particular release	Specify new, customized reference set formats to be used in an Extension